1-1-1980

The use of a reciprocal peer interaction as a means of evaluating a preschool mainstreamed classroom.

Leslie Ellen Weidenman
University of Massachusetts Amherst

THE USE OF A RECIPROCAL PEER INTERACTION AS A MEANS OF EVALUATING A PRESCHOOL MAINSTREAMED CLASSROOM

A Dissertation Presented

By

LESLIE ELLEN WEIDENMAN

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 1980

Department of Psychology

THE USE OF A RECIPROCAL PEER INTERACTION AS A MEANS OF EVALUATING A PRESCHOOL MAINSTREAMED CLASSROOM

A Dissertation Presented

By

LESLIE ELLEN WEIDENMAN

Approved as to style and content by:

Beth Sulzer-Azaroff, Chairperson of Committee

Robert S. Feldman, Member

Roger D. Frant, Member

Melinda Novak, Member

Bonnie Strickland, Chairperson
Department of Psychology

To My Parents

Acknowledgements

I would like to thank the following people for their assistance in the completion of this project.

---The delightful students of the Early Education Center who served as subjects and made this project so enjoyable.

---Judith Souweine, Sheila Crimmins, and all other staff members of the Early Education Center for their cooperation, assistance, and exceptional preschool program.

---The parents of the children for allowing their sons and daughters to participate in the project.

---Glenn Parkinson, my research assistant, for his assistance in data collection and analysis.

---Bob Feldman, Melinda Novak, and Roger Frant, my committee members for their expertise and interest in this project.

---Beth Sulzer-Azaroff, my advisor, committee chairperson, teacher and friend for her undivided attention, advice, support, and encouragement.

---The Bureau of Education for the Handicapped, Department of Health, Education and Welfare, Grant #451AH60923, which funded my graduate training on this research.

ABSTRACT

The Use of a Reciprocal Peer Interaction as a Means of Evaluating a Preschool Mainstreamed Classroom

(Ph.D. University of Massachusetts, 1980)

Leslie E. Weidenmann

A.B. Mount Holyoke College, 1973

M.S. University of Massachusetts, 1978

Directed by: Professor Beth Sulzer-Azaroff

Implementation of P. L. 94-142, the Education for All Handicapped Act, has meant a rapid growth in mainstreaming--the integration of the special nedds child into the regular education classroom. Included in the law is a provision for public education services for special needs preschool children. Many of these children receive services outlined in individual educational plans in integrated preschool classrooms. The rationale behind integrating handicapped and nonhandicapped preschoolers was recently outlined by Bricker (1978) and included social-ethical, legal-legislative, and psychological-educational arguments. Although, the federal mandate includes an evaluation component, such evaluation usually is centered on the accomplishment of individual program objectives. Assessing the impact of mainstreaming on the social ecology of the classroom does not necessarily have to be a part of the evaluation process. The term

mainstreaming can be used to describe a variety of class sizes, types of handicaps, ratios of special and nonspecial needs students, etc. Because of this diversity it is difficult, from a methodological standpoint, to assess the effects of mainstreaming other than on an individual classroom basis.

The present investigation was concerned with the effects of mainstreaming on the social behaivor of special needs preschool children. Behaviors of interest were divided into three main categories: vocalizations, social interactions, and play behavior. To assess the effects of integration, a structured interaction, titled the Reciprocal Skill Development Game, was devised in which handicapped and non-handicapped children were paired together for a brief activity. The intervention was designed as a board game, similar to Candy Land, a popular children's game and employed reciprocal interactions and focused on language skills. Both ingredients, reciprocity and the ability to communicate have been proposed as elements of spontaneous social interactions (Strain & Shores, 1977). The term reciprocal indicates that both participants had equal roles in the interaction unlike more traditional peer tutoring systems where one child serves as a tutor and the other is the designated learner.

Subjects in the study were two special needs preschoolers enrolled in an integrated classroom. A total of 15 children, 8 special needs and 7 nonspecial needs, were in the class. Subject A was a young girl who exhibited a very low frequency of vocalizations and social interactions. Subject B was a young boy with Down's syndrome who also had a deficit in language skills. On standard tests of intelligence, Subject A scored within the normal range, wheras Subject B scored within the retarded range.

Observations were conducted during free-play periods by trained observers. Eighteen behaviors comprising the three major classes were recorded for the two target subjects. A multiple-baseline across subjects design was used to introduce the intervention. Observations continued during free-play for the two subjects. In addition, data were gathered on the nonspecial needs children.

Intervention sessions involved pairing one of the target children with a nonspecial needs student and playing the skill development game. Game periods lasted approximately 15 minutes and were conducted during teacher-directed activity time, the class period preceding free-play. All intervention sessions were supervised by the experimenter and tape recorded. Briefly, the intervention required that

the children take turns playing the game. Steps involved selecting a card from an opponent's hand, describing the scene depicted on the card, spinning a spinner if description was acceptable, and moving a game piece or "man" to the color space designated by the spinner. Social praise as well as tangible reinforcers in the form of stickers were presented intermittently upon correct performance. Immediately following the end of the game, children returned to the classroom for free-play activities.

Data for the two subjects were plotted individually for each of the behaviors recorded. Results were plotted as percentages of occurrence. Data on Subject A showed positive effects. For example, in the category of play behavior, cooperative play increased from a baseline mean of 8% of intervals to a mean of 42% during intervention periods. This was the most dramatic observed. In the area of vocalizations, Subject A showed increased rates of initiations to both peers and adults during intervention, however, the changes were slight. In the category social interactions, only one behavior, affection, showed any marked change. Percentages increased from a baseline rate of 0% to 3% of intervals during intervention. Data obtained from Subject B, revealed an opposite trend. Play behavior

indicated an increase in solitary play during intervention and a slight decrease in cooperative play. Coincident with those data was a slight decrease in the percentage of initiated vocalizations. There were not significant changes in social interactions initiated by Subject B. The data from both subjects were then compared to the mean percentages of intervals obtained from a summary of the data collected on the non-special needs children. This revealed that Subject A's performance during intervention more closely resembled her nonhandicapped peers than it had during intervention. Both Subjects exhibited deficient repertoires during the baseline period in comparison to the nonhandicapped children.

With only two target subjects and opposite trends in the data, a functional relation between intervention and social behaviors could not be established definitively. However, the data are interpreted as indicators of the importance in considering entering repertoires when planning program activities to facilitate mainstreaming. Subject A and B differed considerably in terms of intellectual development and entering skills, even though their handicaps were functionally similar. The results are also discussed in terms of their relevance in evaluating the impact of integration on special needs children.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

C H A P T E R I

MAINSTREAMING FROM AN HISTORICAL PERSPECTIVE

Introduction

The practice of integrating handicapped children into regular education classromms has only recently been widely regarded as a viable means of educating special needs children. Mainstreaming, as this practice is called, has developed in the context of the current movement towards "normalization" of the institutionalized retarded. The principle of normalization has directed the deinstitutionalization movement and resulted in the creation of alternative placements such as group homes, halfway houses, and foster home arrangements for the retarded. In the realm of educating mainstreaming is returning the special needs child to the regular classroom and eliminating segregated education classes from the recesses of public schools.

In addition to the influence of normalization on the development of mainstreaming, a direct influence on its growth can be traced to the research conducted in the 1950's and 1960's evaluating the efficacy of existing special education programs. At that time, considerable efforts as well as financial support had been provided to train teachers and develop special programs to meet

the needs of the handicapped children. Unfortunately, the evaluations failed to produce results that indicated that the children in the special classes improved significantly more than special children in regular classes (Blatt, 1958; Carroll, 1967; Elenbogen, 1957; Johnson, 1950).

The growing dissatisfaction with traditional special education programs increased the popularity for alternatives such as integrating handicapped and normal children. In 1975, passage of Public Law 94-142, the Education of All Handicapped Children Act, guaranteed a free public education for all handicapped children that is specific to their own unique needs and delivered in the least restrictive environment. This law provided the framework for the implementation of special services within an integrated setting.

History of Special Education

The passage of Federal Law 94-142 was a major milestone in the fight for equal rights for the handicapped. It represents the culmination of an ongoing drive by educators, advocates, parents, agencies and legislators alike to provide a procedure for the appropriate education for retarded and other handicapped children. Although, much of the momentum for the movement was generated in the past de-

cade, the origins of special education and the awareness of its need can be found in the early 1800's. When the events leading to P.L. 94-142 are viewed from an historical perspective, the accomplishment of a comprehensive education law takes on an even greater significance. The following sections of the paper will briefly outline some of the more important events and pervasive philosophies that shaped special education.

The first schools for the handicapped were founded in France, under the direction of Itard and his student Seguin, who was an influential writer and member of a group "advocating the rise of a French republic founded upon the principle of 'the greatest good to the greatest number'" (Blanton, 1975). Seguin and Itard were committed to a method of training that was based on the principle of sensory stimulation. Seguin himself credits Jacob Pereire, a Spanish teacher of the deaf and dumb, with the development of the physiological method of sensory training. Adherents to the sensationist philosophy primarily believed that the environment played a major role in shaping one's intelligence. Further, they held that mental deficiency was the result of brain atrophy caused by disuse and lack of stimulation. Therefore, their approach to treatment was to provide sensory stimulation that would awaken the dormant brain and improve the condition of "idiocy." Although

Itard failed to prove his Sensationist theory with his famous case of the "Wild Boy," he and Seguin established a school for idiots in the Saltpetriere in 1838, based on sensory training (Rosen, Clark & Kivitz, 1976). Itard interpreted his failure with the "Wild Boy" to the boy's limited intellect rather than to the training method. Seguin improved upon Itard's original procedure of endless presentations of sensory stimuli by presenting graduated series of stimuli, and achieved considerable success. Word of Seguin's method and his accomplishments spread and by 1846 schools employing his method were opened in other European countries including Germany, but not yet in the United States. In addition, this method provided the basis of more formal educational strategies such as the Montessori method, which was developed at the turn of the century.

In the United States, an increasing awareness of the deviant had developed with the opening of the Worcester State Hospital in 1833. Although schools had not yet been formed, the year 1846 was significant as it was during this year that the Massachusetts legislature established the first commission to investigate the needs and conditions of "idiots." The commission's report emphasized the necessity for providing instruction for the retarded with the result that in 1848 an experimental school was established in South Boston. Seguin's reputation had reached the United

States and all instructors for the first school were sent to Paris to learn Seguin's techniques. Growth of the school led to its move to Waverly, Mass., shortly thereafter. During the following years a rapid increase in the availability of facilities for care and education of the retarded was seen. By the 1880's, eleven states had institutions or training schools for their handicapped populations including: New York, Iowa, Minnesota, Indiana, California, Michigan, Maryland, and Nebraska.

Although the creation of new facilities allowed for the provision of special services for the retarded, subtle changes in the type of service were beginning to be seen. This was partly a function of advancements in other fields, particularly medical science and genetics, and an increase in the number of clients being served by the state's institutions. White and Wolfensberger (1969) described the change as a shift in the general attitude towards the handicapped. They capsulized the period of 1850-1880 as one in which society was attempting to "make the deviant undeviant." This attitude gradually shifted to one in which "the deviant were sheltered from society" (between 1870 - 1890). These changes in the public's perception of the retarded reflect the scientific philosophies and advancements of the times. Seguin and his followers' belief in sensationism was indeed

a belief in the curability of idiocy. It also seemed to assume a unilateral cause for idiocy - that of brain atrophy. By 1866, idiocy was no longer regarded as a "unitary and homogeneous" condition. Down had described the "mongolian type of idiocy" and shortly thereafter, William Ireland published a classification of 12 types of mental deficiency (Rosen, Clark & Kivitz, 1976).

The medical advances in classification of mental deficiency and the increase in popularity of Mendelian genetics and Darwin's theory of evolution altered not only the public attitude towards the handicapped but caused a more general shift in the reigning philosophy towards nativism. Proponents of nativism, including Pinel and Binet, believed in the influence of heredity in determining intelligence and in the irreversibility of neurological defects. Although Binet was a firm believer in nativism, he also advocated education for the retarded. Unfortunately, the general outlook towards the retarded within the frame of reference of nativism was one of protection and care rather than education and rehabilitation. Developing out of these advancements came the Mental Testing movement and the Eugenics movement, both of which had an impact on the type of education made available for the mentally retarded. Before discussing these factors further, it is important to

take a look at what was going on in the realm of Public Education.

The Compulsory Education Act passed in 1876 was intended to provide a framework within which all children would be educated. However, by the 1880's, it became clear that not all children would be able to benefit from the standard educational program provided. As Blanton points out in his chapter on the History of Classification of Mental Retardation, a new problem arose with children who were below average intelligence but were not defective enough to certified as idiots requiring institutional care under the Idiots Act of 1886 (Blanton, 1975). From the educational perspective, then, it appeared that there was an increasing awareness of different degrees of handicaps and that special classes might be able to provide education for some of the intermediate levels. In 1896, a committee investigating the frequency of "subnormal" intelligence reported that approximately 1% of the elementary school population fell within the level of certifiable idiot and those considered ordinary dullard (Blanton, 1975). Further, this committee recommended that special classes be given for these students. Special classes were not authorized by law until 1899, with the passage of the Defective and Epiletic Children Act, more than 20 years after the problem

educating the handicapped first surfaced. And, it was not until 1914 that school systems were obligated to establish facilities to accommodate the children who fell in the range of what would now be moderate retardation.

Returning to the subject of mental testing, it is important to note the relationship between recognition of the need for special education and the growth of mental testing. As it became essential to identify children who were in need of special programs, the sophistication of intelligence testing devices increased. Jacobs and Galton, two pioneers in mental testing, began using memory as an indication of mental capacity as early as 1887. Binet and Simon, originators of the Binet-Simon Scales, included other more complex mental functions such as reasoning and judgement as measures. Binet began publishing his work on mental testing in 1895 although it was not until 1905 that the first Binet-Simon scales were published.

Goddard, in 1910, working at a training school in Vineland, New Jersey was the first to administer the Binet-Simon scales in a widespread manner and analyze the relationship between teacher reports and mental age measurements. After testing 400 youngsters at the school he reported an "amazing correspondence" between groupings based on the test score and those based on teacher reports. As a result, he

suggested that the mental age he used as a means of classification as follows: M.A. less than 2 - idiots; from 3 to 7 = imbeciles; from 8 to 12 = morons; and above 12 - normal. With Goddard's findings it seems that the way was paved for the belief in the stability of the Mental Age score and later the I.Q. It is interesting that such a use of the Mental Test score was not one that was envisioned by Binet. He had recommended that the test results be used as an index of a person's abilities relative to another and not as an absolute measure of their intelligence or ability.

Along with the rapid growth of the mental testing movement was an increase in the popularity of Eugenics which also originated from the Nativist point of view. The single most influential factor in the start of the Eugenics movement, though, was an article by Spencer in 1876 which applied Darwin's principles, particularly survival of the fittest, to the evolution of a social system. Spencer believed that naturally operating forces would favor the growth of the able members of society and not the unfit. As Blanton notes in his chapter on the History of Classification of Mental Retardation (1975), Spencer's philosophy appealed to the wealthy class of post Civil War America which was "fiercely competitive and highly oriented towards achievement" (p. 174). The main support for the Eugenics movement came from men like Andrew Carnegie, John D.

Rockefeller, and in the intellectual sphere, William Graham Sumner, a sociologist from Yale. It was Sumner's belief that inequality among social classes was necessary for society to continue to evolve. Using Darwin's assumption that adapted characteristics could be inherited, he concluded that a progressive society should not attempt to preserve the unfit.

The Eugenics movement, which originated in 1876, carried Sumner's ideas one step further in reaction to the realization that the poor and mentally incompetent were producing numerous offspring. Eugenicists believed in active intervention of such means as sterilization to limit the reproducing capacity of the feebleminded. Although the movement began in 1876, it was not until 1907 that the Eugenics Education Society was founded. During this period, the early years of the 20th century, the supporters of Eugenics were very active in spreading their ideas. For example, in 1910, Davenport established a Eugenics record office, a Eugenics handbook entitled The Trait Book listing inherited characteristics (including feeblemindedness) was published, and in 1914 Goddard's text Feeblemindedness appeared in print and documented the belief that mental retardation was inherited.

The Eugenics movement received a considerable boost

in its support with the publication of a unique study of the family tree of a revolutionary soldier named Martin Kallikak. In his book, The Kallikak Family, Goddard (1912) compared the offspring of Martin Kallikak with his Quaker wife to the descendants of his liaison with a feebleminded girl. It was discovered that of the 480 descendants of the feebleminded girl, 143 were feebleminded along with frequent incidents of alcoholism and illegitimacy. In contrast, the descendants from his marriage had a high incidence of professional people and well respected citizens. The startling differences were interpreted as conclusive evidence of the heritability of mental deficiency. However, the author failed to compare the environments in which the offspring of the two groups of descendants were raised, thus limiting the validity of the data. Nevertheless, these data were cited vigorously to further the Eugenics cause.

The strength of the movement which reached its peak in the 1920's was reflected in the various legal decisions being made as well as in the practices of administrators of state institutions. For example, Barr proposed a statute to the Pennsylvania legislature that would require all retarded citizens to be sterilized. Fortunately, such a measure never gained enough support for passage. However, this did not prevent hundreds of sterilizations

from occurring as a routine matter in state institutions.

Despite the pervasive negative feelings fostered by the Eugenics movement which resulted in generalized fear of the handicapped and mentally incompetent and a separation of these populations from the rest of society, this same period saw the beginnings of a positive trend in educational services (Rosen, Clark, & Kivitz, 1976). Rosen et al. (1976) note that as early as 1917, Wallin advocated that the mentally impaired be educated in special classes designed to meet their individual needs. Improvements in the area of psycho-educational assessment and educational diagnosis were being accomplished through the work of psychologists and educators like Bayley, Thorndike, and Dewey. These developments coincided with the Educational Act of 1921 which enabled the public schools to assign children to special education facilities based on educational incapacity. The trend towards providing special education services grew and by the mid-twentieth century special education was commonplace across the United States.

The growth of a more positive attitude towards the retarded and their need for education was spurred by the rapid decline of the Eugenics movement in the 1930's. Blanton (1975) cites H. J. Muller, a geneticist, as the person primarily responsible for the decline. It was Muller who convincingly put an end to the notion that the

retarded represented an inferior race of people. He stated that "there are no 'pure races' but only populations with variable gene frequencies" (p. 181). His arguments seemed to squelch the raging debate over the need for genetic control of the feebleminded.

The increase in positive public sentiment accompanied both qualitative and quantitative improvements in curriculum materials for special education programs. In addition, attention was being given to training special teachers for the handicapped populations. The mode of special education that emerged during this period was the special class in which trained teachers worked with small groups of handicapped students on a specially designed curriculum. These self-contained classes later (in the 1950's) became the object of a raging controversy in special education, but until that time special education was provided with the assumption that the special class was the best possible method of training the handicapped. The question of the effect of continued segregation of the retarded on self-concept and peer relationships did not come into consideration until recently.

The present trend towards returning special needs children into the educational mainstream has been mandated by federal legislation since 1975. The decades preceeding the legislative landmark saw major changes in public con-

ceptions towards special education and the rights of the retarded and other handicapped groups. The increase in evaluation research during this period was responsible for a large part of the change. The following sections of this paper will examine more closely the relevant social, educational, and behavioral factors that were investigated as variables that had an impact of the successful development of the handicapped students.

The Efficacy of Special Education

Public awareness of the need to educate mentally retarded and handicapped children resulted in the rapid growth of special education during the forties and fifties. At that time, the typical special education format consisted of the separate classroom equipped with special materials and curriculum, headed by specially trained teachers. The rationale for such programming was that instruction could then be geared to the abilities of the children and progress could occur at a rate commensurate with their skills. Not only was this arrangement considered the best for the special education students, but it was believed to be equally advantageous for normal children. Each group could obtain an adequate education that was appropriate for their needs. In addition, the disparate rates of learning in the

two groups would in no way interfere with each group's attainment of knowledge. Consequently, enormous investments in time, money, and effort were made in the development of programs, training of teachers, and manufacture of special education materials. On the surface, this model of special education appeared sound and successful results were anticipated. In practice, however, several commonalities among the nation's special education programs may have ultimately contributed to an ineffective outcome. For example, the majority of special classes were populated by a diverse group of handicapped, retarded, and even emotionally disturbed children who could not be taught effectively or easily by one teacher. Another problem was that many special classes were isolated from normal classes containing the nonhandicapped peer group of the special children. The fact that the peers were segregated undoubtedly limited the degree of contact that occurred between the groups which, in turn, may have influenced the attitudes and expectations of normal peers as well as the regular classroom teachers in a negative direction.

A natural consequence of growth in education is an increasing interest in evaluation of the newly devised programs. Special education was no exception and early on, researchers began investigating its effectiveness from a

number of different viewpoints. The body of research that grew out of the question of effectiveness, portions of which are referred to as the "Efficacy Studies" played an important role in the eventual shift from segregated classes to mainstreamed education.

Although a considerable amount of research was generated on the topic of the value of special class placement, it wasn't until 1968 with the publication of Dunn's controversial article "Special Education for the Mildly Retarded - Is Much of it Justifiable" (Dunn, 1968) that public reaction increased to a level where change was inevitable. This particular article appeared to be highly influential and was widely cited. Its appeal may have originated in the author's documented yet emotional presentation of his position advocating change in the delivery of special education services. Additionally, the positive reaction may have developed because Dunn provided realistic suggestions for making changes and improving the quality of services for handicapped students. At face value, Dunn's suggestions would appear to permit a more effective and efficient system for meeting the needs of special students.

In the following portions of this paper, the issues raised by Dunn (1968) indicating the necessity of implementing changes in the special education structure will be examined. Also, the contributions of individual research

studies to the support of Dunn's arguments will be evaluated for their methodological soundness and the validity of their conclusions.

Necessity for Change.

One of the major purposes of Dunn's report was to demonstrate the need for change. To do so, he outlined four major issues that supported his premise that the current format of special education was unsuited for its intended purpose and, therefore, should be altered. His four points included: the negative effects of homogeneous groupings, the aversize consequences of labeling, the poor results of the efficacy research, and the general improvements in the technology of education.

Homogeneous Groupings.

The first of Dunn's four issues, homogeneous groupings, bespeaks the problems encountered when various elements of the population are segregated from the rest of society. It is Dunn's belief, and many would agree, that homogeneous grouping of students puts the slow learner and the underprivileged child at a severe disadvantage. To support this notion, Dunn cites as evidence, both legal decisions and evaluation research findings demonstrating that underprivileged and lower class black

children perform better academically when integrated with white middle class children or enrolled in schools not operating on a "tracking" system (Coleman, 1966; Wright Decision, 1967). Dunn develops his position further by making an analogy to the educational situation of special needs students. Wouldn't they, like the culturally disadvantaged or racially segregated groups, do better when integrated in a setting with normal peers? He does go on, however, to point out in a cautionary tone, that when the practice of tracking was abolished in Washington, D. C. and the least able children from the lowest track were placed in the normal classroom, the arrangement proved unsatisfactory according to teachers and parents. In retrospect, it seems likely that the dissatisfaction developed due to inadequate teacher preparation and training on how to instruct a diverse class of students in an effective manner.

As a result of the ruling and the dissatisfaction with total integration without the accompanying special training, special education programs developed that were neither a "track" nor a self-contained special class. Instead, the resource room and itinerant teacher systems emerged which seems to combine both integration and special services in a manner that would benefit the students and teachers involved. Dunn concludes, that in the face of

additional court rulings against the legality of self-contained classes for the handicapped, educators should strive to develop new teaching methods that would capitalize on the structure of partial integration and individualized programming. By doing so, educators would be taking a practical step towards the eventuality that special classes will no longer be permitted.

The issue of the legality of special classes seems to confuse the issue of homogeneous groupings and the research questions posed by educational investigators. With the trend towards integration the question of the effects of homogeneous groups, both advantageous and deleterious, seems to have been bypassed in favor of questions concerning the effectiveness of integrated programs. Perhaps the shift in the research focus occurred as a result of the legal decisions affecting educational programming as well as the methodological problems associated with researching the effectiveness of homogeneous groups. Such an enormous issue would be extremely difficult to evaluate effectively. And so, today, 10 years after Dunn first published his article, the factors that significantly affect the success of a special program offered either in an integrated or self-contained classroom have not been fully determined. The problem of factoring out the variables that might be important in any given teaching situation (e.g.,

number of students in class, extent and type of training the teacher received, the type of handicapped students present, age of the students, etc.) are so great that one can only speculate on what might be effective in a particular circumstance. Therefore, it appears that Dunn's first argument against the self-contained special education class could be viewed as a practical and political one. Despite the rationale behind the shift towards integrated special programs being primarily political, numerous research reports, often of poor quality, are cited as evidence that homogeneous groupings are detrimental. As will be seen with many other issues in this area, once a statement has been repeated frequently in public settings, it seems to carry a good deal of importance regardless of the source and validity of the information.

Efficacy Studies.

The second major issue raised by Dunn concerns the results of the "Efficacy Studies" which attempted to compare the effectiveness of special classes with regular education programs. To do so, researchers examined a number of factors such as academic achievement, "personality" development, social adjustment, vocational goals, and number of friends of the special needs students. Generally, the data from special classes were compared to data obtained from special

children enrolled in regular education classes. In many instances the results were conflicting or, at best, unclear. However, there were enough inconclusive findings and a startling lack of positive ones to arouse the concern of educators. By 1968, the general consensus seemed to be that special education was not turning out to be worth the time, effort, and capital that had been invested. The following section will examine, more closely the specific studies involved.

As early as 1932, researchers were interested in the differences between educating handicapped children in the regular classroom and providing special classes. Bennett (1932) and Pertsch (Note 3) compared children in special classes with dull-normal children in regular education classes. Bennett reported that the students in regular classes scored significantly higher on school achievement tests than the students in special classes. Pertsch found similar results in his study which was somewhat more rigorous and methodologically sound than Bennett's. His subject groups were matched according to I.Q. score, chronological age and mental age. He measured motor and manual skills, and personality development in addition to academic achievement. The children placed in the regular classes scores significantly higher on achievement tests than the

special class students and scores on the personality development scale showed boys in the regular classes at significantly higher levels. No differences were found among the girls. In the area of motor skills, no differences were observed despite the fact that the special students received considerably more training than regular students.

With these early studies, evidence began to accumulate suggesting that special education was no better than regular education for the handicapped. However, methodological weaknesses present in these as well as other studies that followed, limit the conclusions that can be drawn concerning the value of special classes for the retarded. The most glaring problem present in the Bennett and Pertsch studies was the lack of control for the academic programs encountered by the two groups of subjects. Because some children were enrolled in a special class, it is conceivable and likely that the curriculum differed markedly from the regular program (otherwise it wouldn't have been "special"). If the programs are different one would expect differences in achievement. This brings up a related concern noted by Johnson (1962) regarding the objectives of special education. It was his position that the areas measured may not have been relevant to the goals of the special education program and therefore not valid indicators of the program's success. It would be similar to teaching a child to add and then

measuring his arithmetic skills with questions on subtraction.

Johnson's discussion of the importance of considering a program's objectives when planning an evaluation brings up two issues that should be major concerns for both program designers and evaluators in any field of study. The first point, simply stated, is that objectives are essential for any program and must be specified by all others who are participating in, observing, or evaluating its operation. It is important that the purpose of a program or set of curriculum materials be understandable to those consumers it is intended to serve. The second point, which is closely related to the first stresses the function of evaluation. Essentially an evaluation should assess how well a particular program has met its specified objectives. Whether or not the original objectives are reasonable or valid for the program is an entirely different issue and one which should be addressed in the planning stages. Johnson (1962) summarizes his view on special education evaluation in the following statement: "Before any meaningful evaluation can be made, the objectives of special education for the mentally handicapped must be defined and the evaluation then made in respect to these objectives" (p. 63).

Efficacy studies continued to examine special educa-

tion in light of its effects on academic performance and social adaptability of the students. Considering that similar dependent variables and research methodologies were used, it is not surprising that the results continued to resemble the findings of Bennett (1932) and Pertsch (1936). A number of the later studies were reviewed by Sparks and Blackman (1965) in an article entitled "What is Special about Special Education Revisited: The Mentally Retarded." Among the studies they cited, Elenbogen (1957) and Cassidy and Stanton (1959) were representative of the methods used and the results found. Both compared achievement and social adjustment and discovered that special students in regular classes performed better academically than special education students. The continued accumulation of data suggesting that special education was not effective from an academic point of view was softened somewhat by the benefits of special education seen in the realm of social adjustment. Both Elenbogen and Cassidy and Stanton's reports supported the placement of children in special education at least for social/emotional reasons. This dichotomy was generally accepted until 1961, when, as Sparks & Blackman point out, Johnson published a study in which social relationships in special classes were examined more closely. She found that children of low intellect held lower social positions in both special and regular education

classes whereas higher intellect children were more highly esteemed. This seems to indicate that no matter where they are placed, special children have lower social positions. This finding further implicates the need for special education.

During this period, it is very likely that special education was not as effective as had been anticipated during its development. However, in its defense, it should be remembered that the validity of the research findings in the early and later efficacy studies was questionable. As mentioned, the studies suffered frequently from a lack of control over potentially important variables such as teacher training, program goals, curriculum materials, student skills and deficits etc. In light of the methodological shortcoming, the criticism leveled against special education may have been unjustified. This does not imply, however, that special education programs couldn't stand improvement. They undoubtedly could have benefited from some changes which would have improved the quality of the educational service to the students. For example, one problem regularly encountered in programs across the country was the diversity of handicapped students receiving special services. Homogeneous grouping really consisted of the segregation of the nonhandicapped from a mixture handi-

capped children. More often than not, the special class appeared to be a "dumping" ground for retarded, physically handicapped, and emotionally disturbed students. The likelihood of success being achieved as measured by group mean scores would not reflect the types of changes that occurred in individual students.

It is unfortunate, that more effort wasn't directed towards determining the necessary components and conditions under which a program would succeed. Instead, the efficacy studies repeatedly demonstrated that the conglomeration called "special education" did not produce significant increases in group mean scores on standardized achievement tests.

The Effects of Labels.

The labeling of a child as mentally retarded, handicapped or special student that has accompanied special education placements was considered by Dunn (1968) to be the third major reason for change in special service delivery. The issue of labeling concerns whether or not the label itself has any demonstrable effect on social adjustment, self-perception, or even academic perfomances. Within the literature the question of labeling effects has been researched from several angles. They include the following areas: (1) teacher expectations of a child's performance as a

function of a label, (2) social acceptance or rejection of a labeled child by peers, (3) attitudes of parents, teachers, and other professionals as influenced by a descriptive label. Each of the areas mentioned will be considered in more detail in the sections below.

Teaching expectations. The question of teacher expectations in relation to special education was an outgrowth of Rosenthal and Jacobsen (1966) research with regular education students. In a series of studies, they demonstrated, although the results have been challenged for their validity, that teacher expectations can become a self-fulfilling prophecy. The labels "rapid learner" and "slow learner" were randomly assigned to groups of children at the beginning of the school year. Teachers were told that specific individuals were expected either to do well or poorly during the school year, depending on their pre-assigned experimental condition. A pre-test, post-test comparison of standardized achievement test scores showed, in many cases, that children labeled "rapid learners" had made greater gains than the designated "slow learners." The extension of this finding to special education assumes that certain expectations of a child's ability accompany the label "mentally retarded." Although the "Rosenthal Effect" as this finding is called, has been severely criticized for methodological reasons, thus raising questions about the validity of

the effect, the notion of the label "mentally retarded" causing changes in teacher behavior seems to have considerable common sense appeal. What is unclear, though, is what type of behavioral changes might take place, and what effect they would have on individual students. It is possible that a teacher might provide more individual instruction, or use different curriculum materials, or plan a host of special activities for a child labeled retarded, all of which might facilitate school achievement. Dunn's reliance on the Rosenthal research was criticized severely by MacMillan (1971). In a somewhat sarcastic comment he expressed his disbelief in the Rosenthal effect.

> If one could extrapolate so easily from the Rosenthal and Jacobsen (1968) work as is implied by Dunn, the problem could be solved immediately by simply labeling the children under consideration "gifted" and thereby increase the teachers/expectancy for them to succeed (p. 6).

Two studies, Soule (1972) and Yoshida & Meyers (1973) examined the expectancy effect with mentally retarded populations. Soule adapted the principles of the Rosenthal studies to severely retarded children living in cottage settings. Teacher bias was induced by randomly assigning children a label indicating they were "expected to show progress in

the near future." Pre- and post-test measures on intelligence and behavioral ratings were taken on the experimental subjects and a control group. The control subjects were matched for age, level of intelligence, and motor skill performance. Comparisons of the pre- and post- scores of the two groups did not demonstrate any effects of the experimentally induced bias. One problem noted by the author was the lack of validation of the existence of bias in the cottage workers. However, this at least indicates that bias may not be as predictable as originally believed.

Yoshida & Meyers (1973) studied teacher expectancy effects with educable mentally retarded children in a public school setting. Both regular and special education teachers were asked to predict the future achievement levels of a child on a concept formation task at four different intervals. The teachers viewed a video tape of the child working on the task and were told that the boy was either a regular sixth grade student or an educable mentally retarded student. The repeated measures, analysis of variance, performed on the data did not reveal any significant differences in the predicted achievement levels. In this instance, expectancy effects resulting from a label were not present. The importance of the results of this study is twofold. Primarily, it casts doubt on the existence of a generalized teacher expectation effect, and secondly it

suggests that other factors may be the more crucial determinants of teacher behavior. As the authors note, the labeling perspective doesn't account for (a) possible changes in expectancy as a result of direct observation of the student's behavior, and (b) potential role differences of the special education and regular classroom teacher. This study, though, like Soule (1972) began with the assumption that the labels would produce a bias. The fact that no real differences in expectancy scores were found suggest that such assumptions may not be warranted.

Social acceptance. The second facet of the labeling issue concerns the social adjustment of the labeled child within his peer group. This is generally measured by a sociometric questionnaire and is reported in terms of the labeled child's social status within the non-labeled peer group. Two early studies examined the status of children labeled "mentally retarded" among regular elementary education students. Johnson (1950) using a sociometric instrument found significantly lower peer acceptance scores for retarded children than for normal children in the same grade. Baldwin (1958) employed the Ohio Social Acceptance Scale, a forced-choice scale on which all children in a class or group rate each other on the degree to which they would want each individual as a friend. Baldwin sampled fourth, fifth, and sixth grade children in a large public

school system and like Johnson (1950), found lower social acceptance scores for the mentally retarded than nonretarded students. Along with the sociometric measure, Baldwin interviewed all teachers and students who participated in the study. From their comments, Baldwin concluded that the major factor contributing to the poor social acceptance of the retarded was the "antisocial" behavior they exhibited.

Later studies examined the labeling question and its effect in different settings and with other populations. For example, Mercer (1971) found that retarded children from the inner city were quite different than suburban retarded children on the dimensions of educational background and physical characteristics. Bruininks, Rynders, and Gross (1974) then compared the Peer Acceptance scores of the mentally retarded children in resource room or regular class placements in both urban and suburban areas. No differences were reported for the class placement factor but slightly higher acceptance ratings were obtained by urban retarded children compared to urban nonretarded children when rated by children of the same sex. Goodman, Gottleib, and Harrison (1972) looked at the effect on social acceptance of both integrated and segregated mildly retarded children attending a nongraded elementary school. Using the Peer Acceptance scale, they found that both integrated and segregated educable mentally retarded children were generally

less accepted than their nonretarded peers. However, the authors suggest that an alternative explanation for the lower degree of acceptability may be that the educable mentally retarded children were bussed in from other neighborhoods. The other children in the school were from the local area.

Gottleib and Budoff (1973) continued the study of labeling effects in the nongraded school setting by looking at differences in acceptance levels of educable mentally retarded children in programs that were housed in schools that differed in physical structure.

Interest in this question grew out of the assumption that the organization of the nongraded school program would foster higher levels of acceptance of educable mentally retarded children than a traditional school. The rationale underlying the assumption is that the individualized program with its flexibility and criterion referenced system of evaluation provides students with a greater awareness and perhaps greater acceptance of individual differences. Gottleib and Budoff (1973) compared the social acceptance of educable mentally retarded children in a no-interior wall nongraded program with those in a traditional "egg-crate" school building. It was reasoned that educable mentally retarded children would be highly visible in the unwalled school thus increasing the opportunities for non-

retarded children to become familiar with them. Based on the findings of Goodman et al. (1972) in which greater rejection was found for educable mentally retarded children in the nongraded school, the authors predicted that the educable mentally retarded children in the unwalled building would be rejected more often. The results supported their hypothesis--educable mentally retarded children in the unwalled school were known more often by their peers but not selected as friends. This finding of lowered acceptance in the nongraded setting by both Goodman et al. (1972) and Gottleib and Budoff (1973)was accounted for by the authors in the following explanation:

> The unexpected greater rejection of integrated children was that they were perceived as nonretarded and were expected to conform to the behavioral standards of other "normal" children (p. 15).

It is interesting that this explanation is offered in studies reporting lower acceptances of children labeled mentally retarded, as it could be interpreted as support for a position favoring the use of labels. It might be argued that the label alters the standards of acceptable behavior for the children to whom the label is applied. Once labeled, perhaps less rigorous social standards are applied. Thus, because a child is called retarded, allowances might

be made for any instances of deviant or extreme behavior.

Accompanying this interpretation, there appears to be an implicit assumption that handicapped children are never "accepted" socially to the degree of a nonhandicapped child. That is, almost be definition, they can never achieve the social status of a normal child. If this is the case, and there are certainly no data that conclusively demonstrate either position, the methodologies used in the acceptance studies are inappropriate. In the studies reviewed the social acceptance of a retarded child was measured in comparison to a normal child. Perhaps it would have been more useful to compare the relative degrees of acceptability of a variety of labeled populations such as physically handicapped, emotionally disturbed, mentally retarded, and learning disabled. Then if differences were found in the degree of acceptance, researchers could begin to focus on the factors that are discriminable and important to the raters, that is, the normal peers.

The bulk of the literature on the effects of using the mentally retarded label was reviewed by MacMillan, Jones, and Aloia (1974). They interpreted the lack of definitive findings from the research as an indication that the investigators had failed to specify clearly what the key issues surrounding the labeling question were. Therefore, they suggested that the methodologies employed were likewise not

suitable for obtaining the kind of data needed to make conclusive statements regarding the effects of labeling. They also point out that the statements denouncing the use of labels that appeared frequently after Dunn's (1968) article was published were somewhat premature and lacked research support.

MacMillan et al. identified three separate issues that seemed to be of concern to investigators studying the labeling question. They distinguished: (1) dissatisfaction with the effectiveness of self-contained classrooms, (2) the reliance on intelligence tests which may be biased for culturally deprived children, for special class assignments and (3) the effects of the label on various outcomes once a child has been labeled. These issues can be broken down even further and the authors list several subtopics. (The reader is referred to the original publication for a more detailed discussion of the problems noted in clarifying research questions as well as selecting research designs to answer those questions.) Although it has not been covered in detail here, it should be emphasized that the contribution of MacMillan et al.'s review is that it begins to untangle the confounding variables present in the studies on labeling and suggest appropriate methods for investigating the question in less complex ways.

One study conducted by Chennault (1967) attempted to improve the social acceptance of unpopular educable mentally retarded children through direct intervention rather than assess the level of acceptance or rejection by peers. Intervention consisted of organized group activities such as the planning, rehearsal and presentation of a single dramatic skit. The shared experience lasted approximately 5 weeks and required two fifteen-minute sessions per week. The determination of an increase in social acceptability was made by comparing pre- and post-test measures on a social acceptance scale administered to the peer group. In addition, the individual perceptions of the educable mentally retarded students were measured before and after the intervention. The results indicated that the experimental subjects improved significantly on both the peer acceptance scale and the self-perception of social status measure. Although it seems likely that the improvement was directly related to the organized group experience the author points out that other related factors may have contributed to the status change. For example, experimenter attention, the concrete reward given to the class at the end of the activity, and removal from the classroom twice a week were suggested as possible factors.

The importance of Chennault's study is not necessarily

limited to the determination of specific factors that will reliably improve the social status of educable mentally retarded children. But, it provides an alternative method for examining the general question of the acceptance or rejection of labeled children. The experimental study examining the effects of a direct intervention also provides data that is potentially more useful to educators faced with the question of how to effectively mainstream a handicapped child into the regular classroom. This study may also serve as a model for other researchers to follow in order to determine the variables that should be considered for effective mainstreaming.

Attitudes. An addition to studying the effects of labeling by measuring social acceptability and expectation effects, investigators have frequently looked at the attitudes of parents and peers towards children labeled mentally retarded as well as the self-perceptions of the labeled children themselves. A common procedure used to assess the influence of the label on attitudes is to administer a semantic differential scale to subjects who have either viewed a video taped sample of a child's behavior or read a description of a particular subject. Depending on the assigned experimental condition, subjects are informed that the given information refers to either a mentally retarded

(or similar label) or normal child. The types of items presented on the semantic differential scale might include the following: self-reliance - dependent; neat - sloppy; employable - unemployable.

This technique was used by Cook and Wollersheim (1976) to examine the effects of the mentally retarded label on the attitudes of seventh and eighth grade students. In addition to the labeling factor the investigators examined the effect of contact among normal peers and mentally retarded students on their perceptions. The authors hypothesized that the labeling effects would vary with peer contact and would differ depending on whether contact was with educable mentally retarded (EMR) or trainable mentally retarded (TMR) students.

A 2 x 3 factorial design was used to compare 3 levels of contact (EMR, TMR, and no contact) with the two labeling conditions (mentally retarded label, no label). Fifty students were chosen from three separate schools encompassing students from similar socio-economic backgrounds. The schools differed in the availability of services for retarded children. One school had a class for EMR students, a second school had one class for TMR and the third school did not provide special services. The type of contact, therefore, being compared included the opportunity to eat

lunch, share recess and music classes, and participate in all school activities together. Measures of perceived behavior and commitment to involvement were taken in addition to the semantic differential scale previously mentioned. The perceived behavior measure presented subjects with both adaptive and maladaptive behaviors on a six-point scale. The subject's task was to indicate how sure they were that the person described in the passage would emit the behaviors. The commitment measure gave the subjects the opportunity to show their willingness to work with mentally retarded students from special education classes.

The results of the investigation which were subjected to an analysis of variance showed a significant main effect for labeling and a labeling X contact interaction on the semantic differential measure and perceived behavior scale. As Cook and Wollersheim summarize, the responses of the control (no label) condition were significantly more positive than the label condition. The contact variable was not found to be significant although it did interact with the labeling condition. The lack of significant effects regarding the contact variable suggests that the type of contact assessed in this study may not necessarily affect new behavior in a uniform way. The potential contact among subjects and retarded children may have been so varied that

the effects were washed out. For example, contact for one subject might have been observation only, whereas another subject may have actually played with a retarded child on the playground. Observation without interaction may not necessarily have a positive effect on perceptions. Neither does the procedure allow for differences in the behavior of the retarded children with whom the subjects had contact. If by chance the children just entered school or were transferred from an institution, their behavior and appearance may be perceived as extremely bizarre or deviant. To fully evaluate the impact of contact with the mentally retarded it seems essential that additional factors such as type of contact,length of contact, the behavior exhibited by the retarded students, etc. be considered.

Differences in the degree of commitment for contact and noncontact groups were found to be significant in Cook and Wollesheim's study, although not in the predicted direction. It had been expected that groups who had contact with the handicapped children would have a more positive commitment than the noncontact group. The data, however, showed that the noncontact peers were significantly more positive and indicated a greater level of commitment than the group who had contact with the TMR children. There were no significant differences in commitment between the EMR contact and no-contact group.

This surprising result raises questions about the nature of the "contact" experienced by the normal peers and whether or not different types of contact have differential effects on a measure of commitment. This particular study may have tapped only one type of "contact" which showed a negative effect on peer commitment. For example, although contact was defined as the opportunity to interact during lunch and recess, it is possible that contact may have been primarily observational. It seems unlikely, that in the absence of specific arrangements, students would initiate frequent social interactions during lunch and recess periods with children from different classes or grades. This might be particularly true if the behavioral and social skills of the classes were at widely different levels. It is also possible that observation of groups with unique or discrepant behaviors without any accompanying interaction or explanation of such behavior might result in negative perceptions as demonstrated by a lesser commitment. This explanation may be one of several that could account for these findings. The lack of specification of types of contact, observed behaviors of the retarded children, degree to which the special classes were integrated into the school, etc. may all be important factors in determining the impact of contact and the mentally retarded label on normal peer

attitudes towards the retarded.

Two experiments by Gottleib (1974, 1975) examined the effects of the mentally retarded label on the attitudes of third-grade students. A video tape of a child actor named John was shown to subjects, half of whom were informed that John was a mentally retarded child. The other half were told that John was a fifth grade student. In addition to labeling effects Gottleib studied the contribution of particular behaviors exhibited by the child to the resultant attitudes.

In the first study (Gottleib, 1974) John displayed academically competent behaviors in the film. A control tape showing John behaving in an incompetent manner was also developed. A 2 x 2 factorial design (label X behavior) was used with random assignment of subjects to each of the four experimental conditions. Subjects either viewed John performing well academically on the task or acting in an incompetent manner. Half of each group believed that John was a retarded child while others believed he was a normal fifth grader. The measures of attitude included a 5-point semantic differential scale and a modification of the Cunningham Social Distance scale, a 6-point measure indicating a subject's agreement with a number of descriptive statements about John. The two scales were administered to

all subjects immediately following the viewing of the video tape. The results showed, interestingly, that the attitudes towards a child exhibiting academically competent behaviors were significantly more positive than an incompetent child regardless of the labeling condition.

In the 1975 study, Gottleib examined the impact of social behavior and labels on student attitudes. John demonstrated acting-out and aggressive behaviors on the experimental video tape and the control tape showed him behaving in a socially appropriate manner. Following the same procedure and design as the 1974 study, Gottleib administered semantic differential and social distance scales to the four groups of students following the tape viewing. As in the earlier study, the attitudes towards John were more favorable when he engaged in socially appropriate behavior, however, a significant main effect for labeling was obtained on the semantic differential scale. Significance was approached on the social distance measure and both scales revealed significant behavior X label interactions. In addition to the analysis of variance, the omega statistic was employed to determine the contributions of the label and behavior on the attitudes. For the semantic differential scale, it was calculated that 20.4% of the variance in the ratings was due to the label, 21.3% from the behavior and 6.4% from

the interaction of the two. On the social distance measure, the effect of the behavior was even more pronounced, accounting for 35.1% of the variance. The label and the interaction, on the other hand, only accounted for 5.7% each of the variance.

The importance of Gottleib's work is that his findings suggest that it is not the label per se that determines a person's attitude but the label within the context of a particular sample of behavior. This point becomes quite clear when the data from both studies are viewed together. In the 1974 study when the label was paired with academic behavior, the label did not produce any significant difference in the attitude measures. The behavior of the subject was the major factor in the formation of attitudes. In the 1975 study, however, the label was paired with social behavior and did have significant effects. But the results of the omega statistic showed that the behavior component accounted for more of the variance in the attitude scores. Gottleib interpreted this difference as an indication of the complexity involved in the effects of labels on attitudes as well as the lack of consistency and predictability of the findings. Gottleib's position is summarized in the following statement:

> It can be concluded from the two studies that labels are important contributors to

> negative attitudes, but only under certain circumstances, for example when they appear in combination with aggressive, acting-out behavior. The two studies also suggest that blanket denunciation of labeling as ubiquitous contributors to negative attitudes are unfounded. (1975, p. 584)

Gottleib goes on to say that a profitable approach to studying labeling effects would be to specify the behaviors and conditions under which the label would have a negative effect on attitudes.

It is interesting that Gottleib's results pointed out the complexity of the labeling effect whereas other studies, such as Cook and Wollersheim (1976), discussed earlier, concluded that labels have a definite negative impact on attitudes. The differences in certainty with which the two studies are interpreted may be a reflection of the difference in procedure and the appropriateness of the methods used. A superficial examination of the two procedures shows striking similarities: both use semantic differential scales to measure attitude, group factorial designs, random assignment of subjects and use of the mentally retarded label. Despite the similarities, a major difference can be seen in the type of information given to the subjects regarding the

labeled child. On the one hand a written description was presented whereas on the other, a videotape of the child in question was shown. When only provided with a written description, it is possible, that subjects rely more on the label than when they are able to observe the child's appearance, social behavior, verbal behavior, etc. It may be a question of basing opinions on the most salient information available. The videotape method more clearly approximates a real-life situation and certainly seems to have a greater degree of face validity. Validity could be improved even further, perhaps, by the use of actual retarded children in the sample video tapes rather than employing an actor. One study to be discussed below, used such a procedure.

Investigations of the impact of labels on attitudes were not limited to the study of peer groups of the labeled populations. Assessments also have been made of the attitudes of school administrators, teachers, parents, graduate students and the labeled children themselves.

Teacher attitudes towards integrating children labeled educable mentally retarded, emotionally distrubed, and learning disabled were examined in a study by Shotel, Iano, and McGettigan (1972). Subjects were teachers from schools that were instituting integrated programs with re-

source rooms. All teachers participated in several in-service training meetings on mainstreaming just prior to the beginning of the school year. The control group consisted of teachers from schools with self-contained special classes. Measures of teacher attitudes were taken at the beginning and end of the school year by asking teachers to respond to a 13-item questionnaire on the placement of handicapped children. The authors obtained a number of interesting results: (a) initially, the experimental teachers expressed a much more positive attitude towards integration of all three classifications of handicaps, (b) the posttest showed a closer correspondence between the attitudes of the two teacher groups. (Although the experimental group's responses were less positive than the pretest, they were still slightly more positive than control subjects' scores.) (c) The attitudes towards EMR children were the least positive of the three groups regardless of the teachers' experience with resource room programs.

The decrease in the experimental subjects' responses to integration at the end of the year has serious implications for mainstreaming programs. It is possible, as the authors suggest, that the initial positive attitudes may have reflected the success of the training meeting on the philosophy of integrating the handicapped which all experimental teachers attended at the beginning of the year. The

posttest attitudes, although less positive, were based on an entire year of observation of the special needs children in the classroom and are probably a more accurate representation of the teachers' attitudes. The low degree of positive responses may have resulted from difficulties experienced in integrating the children socially or slower than expected academic progress. The authors conclude:

> It remains an open question whether or not, at the elementary school level, programs providing increased integration of educable mentally retarded children can be successful in schools using the conventional grade organizational pattern. (p. 683)

They also suggest the importance of providing support services for teachers in mainstreamed classrooms, in the form of periodic inservice training, opportunity to observe the resource room in operation, and frequent communication between the classroom teacher and resource room staff. In addition, it would seem important to provide adequate instructional materials and train teachers with specific management skills.

The chance of success of an integrated education program for handicapped children will undoubtedly be affected

by the attitudes of the teachers who will be asked to implement the procedures. Teachers, though, are not the only members of the school community who might influence the outcome of a mainstreaming program. Payne and Murray (1974) considered that the attitudes of school building principals would have an impact on how well specific programs are carried out. To examine the positions taken by principals on the issue of school placement of handicapped children, the authors distributed a questionnaire on mainstreaming to 50 urban and suburban elementary school principals. The questionnaire surveyed the principal's willingness to integrate handicapped children, their feelings on the type of handicap their school could service, and resources that would be necessary to implement a mainstreaming program. With 65% and 70% of the urban and suburban principals respectively responding, the data indicated that suburban principals were significantly more supportive of integration than urban principals. Although the degree of support differed, the trend of acceptance across the seven different categories of handicaps presented in the questionnaire were very similar. Both urban and suburban principals responded much more positively towards visual and auditory handicaps than towards two levels of mental retardation, EMR and TMR.

In terms of the services needed, both urban and suburban principals ranked inservice training as the most important need for teachers, followed by resource teachers and resource room programs.

Despite the small sample of principals participating in the study, the data suggest that integrated programs would have a better chance of success in a suburban setting due to greater support from the school administration than in an urban setting. Unfortunately, the study does not provide an indication of what factors affect the principals' attitudes and thus do not benefit those who are faced with the task of implementing mainstreaming programs in urban settings.

The effects of labels on attitudes of mothers and graduate students were studied by Seitz and Geske (1976) using videotaped samples of mother-child interactions. In addition to assessing the labeling effects, the investigators were interested in examining the way in which retarded and nonretarded children interact with their mothers and whether or not differences in behavior could be perceived by observers. The observers consisted of the mothers of both the retarded and nonretarded children and a group of graduate trainees in clinical psychology. All subjects viewed three 10-minute segments of videotaped

interactions of mother-child pairs in an unstructured free play setting. One tape showed a labeled retarded child with his mother, one showed an unidentified retarded child-mother pair and the third was a nonretarded child-mother dyad. The order of presentation of the tapes was counterbalanced across four experimental groups. Immediately after viewing each tape the subjects were instructed to fill out a 10-point social distance scale for the child in the film and to rate each child on several personal characteristics such as happiness, appeal, curiosity, etc.

The results of the study showed that both mothers and graduate students can discriminate retarded children from nonretarded children on the basis of behavioral observation. It was not necessary for subjects to rely on the label to differentiate the two groups of children. The authors also reported that the quality of the interactions of mothers and their children were different as a function of whether the child was normal or retarded. They found mothers of retarded children to be more controlling in a play situation than mothers of nonretarded children. This raises the possibility that the subjects discriminated the two classes of children on the basis of the mother's behavior rather than on the specific repertoires of the children. Unfortunately, this question cannot be answered by the data.

The effects of the label was very pronounced in the interpersonal attraction ratings obtained from both the mothers and the graduate trainees. The retarded child who was not so identified in the tapes was consistently rated as less happy, curious, independent and appealing than when the label was provided. This finding, the authors felt, was supportive of MacMillan's (1974) notion that the label has a "protective function eliciting altruistic attitudes" towards the retarded child. The data seem to indicate that the label itself provides the observer with an explanation or account of any deviant behavior that might have been noticed.

The protective function of the mentally retarded label is considered by some to be a positive effect of the labeling process. However, those who favor the abolition of classificatory labels argue that the negative effects and the labeled individuals such as poor school achievement and a low self-esteem far outweigh any beneficial effects. The research investigating the topic area of self-concept of the labeled child, in actuality seems to have examined the effects of special class placement or resource room programming on the children rather than the effects of the mentally retarded label itself. From the data gathered in the placement studies, it appears that authors may have

interpreted the findings as indications that labeling effects existed. For example, Meyerowitz (1962) found that elementary school children's self-concept dropped following special class placement. Jones (1972) similarly found lower self-esteem in retarded high school students who attended special classes than nonretarded students in the regular classes. Such findings have been cited as evidence of the stigma associated with labels. It would be more accurate to say that whatever stigma there is, it is probably in part associated with attending special classes and perhaps being segregated from the nonretarded students.

Carroll (1967) examined the effects of two types of school programs, segregated and partially integrated, on both self-concept and academic achievement of educable mentally retarded students. To measure self-concept, Carroll employed the Illinois Index of Self-Derogation (IISD) a scale that standardized on EMR children in Illinois. Items on the index require the subjects to select one of two sentences that most closely represent themselves. One of the two sentences describes a socially desirable attribute while the other presents an undesirable description. A pre-test - post-test design was used and subjects were asked to respond to the index at

the beginning and end of the academic year. The results indicated that EMR children in a segregated setting were significantly more derogatory than EMR children in a partially integrated program. Both groups of EMR students showed significant improvements in reading, spelling, and arithmetic during the academic year. The children in the partially integrated program, however, made significantly greater gains in reading than the segregaged students.

The achievement data as well as the self-concept results of this study support the concept of mainstreaming special needs children into the regular classroom. It appears that the partial integration program leads to a higher self-concept than a totally segregated special education class. However, these results should be interpreted with some caution as they do not indicate whether or not the retarded children are still somewhat stigmatized by their involvement in the partially integrated program. For example, if other measures were used such as peer acceptance scales etc., it might be discovered that the EMR children are still being rejected by their normal peers, because it has been shown that retarded children are generally less well accepted than their normal peers in an integrated setting (Gottleib & Budoff, 1973; Goodman et al.,

1972). Neither do the data specifically enable researchers to conclude anything concerning the effects of labels on perceived stigma or self-concept.

Improvements in Educational Technology.

The fourth and final topic included in Dunn's (1968) original paper outlining the reasons for changing the then pervasive model of special education, the self-contained special class, was the improvement in available teaching methods. Dunn recognized that the current practices in educational programming in 1968 were considerably better than what had been used in the early fifties when special education was expanding rapidly in the direction of segregation. Innovations resulting from the work of educators, psychologists and learning theorists could be found in the area of curriculum development, teacher training, structural improvements in the physical plant, and the use of computer technology. As the newer developments became commonplace in many regular educational systems, Dunn noted that the regular classroom setting was better equipped to service handicapped youngsters; thus eliminating, in part, the original justification for the self-contained special class.

In the field of curriculum development, Dunn pointed to the growth of individualized instruction as a major im-

provement making the regular classroom more amenable to the special student. Teachers already using individualized programs with their students would not have to make much of an adjustment to incorporate a handicapped child with his own tailor made program. These teachers are continually faced with students working at different levels of proficiency proceeding at different rates. Thus, the handicapped child would not present any additional problem as long as the appropriate set of materials had been prepared.

Since 1968, other developments in the area of teaching methods and curriculum development have increased further the feasibility of successfully integrating handicapped students. For example, from the work of behavior analysts has come the personalized system of instruction (PSI) (Keller & Sherman, 1974), peer tutoring systems, skill development training packages for the retarded, principles of behavior for classroom management, competency based systems, and the use of behavioral objectives. Reports of new developments, many of them found in journals such as Behavior Modification and the Journal of Applied Behavior Analysis, demonstrate clearly how the use of these procedures can result in improved performances with a variety of student populations.

To accompany the improvements developed in the laboratories and field settings, methods must also be avail-

able for transmitting the new information to aspiring teachers and those already in service. Indications of the impact of a new technique should then be reflected in both teacher training programs and the classroom. Already such topics as programmed learning, and behavioral objectives have become standard fare in Educational Psychology texts (Anderson & Faust, 1975; Gage & Berliner, 1977) and it is hoped that information about other successful models of teaching and technological innovations will be added in the future.

Aside from the new developments in teaching methods and curriculum design, Dunn (1968) noted that the organizational structure of the classroom itself underwent a transformation in many schools. No longer was the traditional system of one teacher per grade assigned to a room of 20 to 30 students seated in desks neatly arranged in rows--the only format. In the older buildings limited to the restrictions of the existing physical plant, innovations in seat arrangements occurred such as the creation of work groups, or stations in place of rows. Team teaching and teacher specialization in certain subject areas has also become more common at the elementary level. For the newly constructed schools, architectual plans reflected the spirit of innovation, showing no-wall interior buildings and

quad designs. Although the traditional school building certainly does not limit the possibilities for effective integration of special needs students, the newer designs coupled with innovative teaching methods seem to allow greater freedom and flexibility in programming thus providing better services for all students.

Recently, several models designed for the integration of mildly retarded students used in eight school districts were discussed by Guerin and Szatlocky (1974). The authors observed that the majority of the programs incorporated a learning resource center in conjunction with integration of the handicapped students in the regular classroom. Most schools operated a system of partial integration in which the special students spent a portion of their day with their normal peers in the regular class and a portion of the time in the resource center where they would receive special individual instruction in the areas they required. The degree of integration seemed to be determined by the degree of the child's handicap and the skills he or she possessed.

Another area of importance in general educational services noted by Dunn (1968) was the increase in the number of ancillary school personnel. The addition of school psychologists, guidance counselors, teacher aides, re-

source room staff, learning specialists, speech pathologists, etc., has theoretically made the job of working with handicapped youngsters easier for the classroom teacher. The availability of support services, for example, reduces the problem of not enough time for preparation of materials frequently cited by teachers as the reason for not wanting special needs children in their classes. Ideally all members of the staff can combine their efforts and plan a program to remediate an individual's school problems. This format has been endorsed by state legislatures with the passage of Special Education Laws such as Massachusetts' Chapter 766. The regulations require that prior to the delivery of special education services the child in question be given a core evaluation by educators, psychologists, speech pathologists, and medical professionals.

One final area of improvement noted (Dunn, 1968) was the field of electronic hardware. This ever expanding area has grown tremendously since Dunn's article first appeared and the technological products of the electronics industry including calculators, small computers, teaching machines, video tape systems, audio equipment, laboratory materials, etc. have become more and more common in public school systems. One of the most appealing features of this equipment is that it can often be used to create self-instruction

centers for students. The value of self-instruction has been discussed at length by many authors. B. F. Skinner, for example, wrote an influential article in 1961 titled "Why We Need Teaching Machines" and outlined a number of reasons they would be an asset to the classroom. Among the most compelling were the following: (1) improved efficiency of both student and teacher use of class time, (2) student has the opportunity to be reinforced frequently, (3) student can receive immediate feedback on his performance, (4) the student can progress through an orderly sequence of material at his own rate.

The improvements in educational technology that have occurred since the conception of the self-contained special class have altered the climate of the regular classroom in terms of the breadth of services it can provide for a variety of students. With the advent of better curriculum materials, the growth of individualized instruction, improved facilities, more efficient use of teacher skills and teaching time, the addition of auxilliary personnel, and the greater availability of sophisticated equipment, the regular classroom is able, theoretically, to instruct effectively both handicapped and nonhandicapped students alike. Today, with the trend towards integration firmly established, the focus of educators must shift to improving the quality of teacher training to keep up with the

rapid developments in the field of teaching methods. Disseminating new information and providing teachers with the necessary management skills required to operate an integrated setting will be important for the success of mainstreaming programs.

Summary

The need for improvement in the quality of special education services has been recognized since the publication of the initial evaluation studies conducted in the 1950's. The results of several of these studies implied that the existing special programs were no better than the education special students would receive if enrolled in the regular classroom. However, this interpretation was tentative as the validity of the research was questionable, yet it did provoke a long debate among educators about the direction in which change should occur and the type of educational model that would best meet the needs of the handicapped students. Although, these questions have not been answered definitively, the controversy brought these issues into the public domain where they were picked up by parents, legislators, and other advocates for the right of the handicapped. The culmination of their efforts was the passage of P. L. 94-142 in 1975 which requires that a free public

education in the least restrictive environment be provided for all handicapped children. The literature reviewed in the preceding sections represents the major issues in the controversy surrounding special education offerings. Although vast amounts of data have been collected by numerous investigators, only a few general principles can be derived from the research. This reflects, in part, the presence of conceptual and methodological problems in the studies. It also suggests that the goal of determining the best method of educating special needs children is one that requires an enormous effort and perhaps may be larger than any one researcher can cover adequately given financial and time limitations.

The first issue to be explored by researchers was the efficacy of special education. This was studied by comparing special students in segregated special classes to those enrolled in regular education programs along the dimensions of academic achievement, social adjustment, and self-concept. The findings of several representative projects suggest, tentatively, that an integrated or mainstreaming program results in greater academic achievement and higher self-concept than a segregated program. The studies of segregated programs, on the other hand, indicated that students were more favorably regarded by their peers than when they were in the regular classroom. That is, they obtained

higher social position ratings from normal peers when they were placed in the segregated setting. One study, however, suggests that the social position of the handicapped student is lower in both settings than the normal student. (These data seem to indicate that when the behavior or responses of the handicapped themselves are examined, the integrated programs prove more advantageous. Although, the overall advantage may be mitigated by the loss in social prestige if peers have less regard for the handicapped when they are enrolled in the same educational program.)

The largest area of research relevant to his particular question concerned the effects of labels such as mentally retarded, learning disabled, or emotionally disturbed which often accompany the receipt of any special services. This question was approached from several vantage points including teacher expectation effects, social adjustments, judgements, attitudes, and self-concept. In most instances, labeling studies did not consider simultaneously, the question of an interaction between the use of a label and the type of special education programs (segregated or mainstreamed) the labeled child received. These findings, therefore, cannot support directly the research evaluating special education models, although, many authors have speculated in this direction.

Whether or not labeling effects were obtained seems to be dependent on the specific questions asked, the measures used, and the definition and connotations of the term "labeling effect" employed by the authors. The research reviewed appeared to fall into three levels or categories according to their findings. The first level is illustrated by the studies that did not obtain any significant differences in their dependent measures as a function of labels. This occurred in the area of teacher expectations and judgements about handicapped children. The lack of labeling effects can be interpreted to mean, perhaps, that variables other than the label were more salient. That is, other aspects may have provided the observers with more relevant information on which to base their judgements.

The second level can be seen in the studies reporting significant differences in their dependent measures, but where questions arise as to whether the differences were solely a result of the label. Some of the research in the area of social adjustment, peer acceptance, and attitudes falls into this category. The results obtained from social distance and peer acceptance questionnaires, filled out by the normal peers, indicated that the special class students were less accepted and held lower social status than their normal peers. It was also found that a

greater rejection of special needs students occurred when they were not identified as such and when they were integrated into non-graded school programs. However, these findings might also have been influenced by other variables such as the fact that in one study (Goodman, et al., 1972) the special needs children all were bussed in from different neighborhoods, in that there were no controls for the behavioral differences in the repertoires of the different children. It is possible that the obtained scores were due to the behavioral discrepancies observed rather than the label itself. Other factors also might have affected the scores, including a lack of familiarity or even differences in socio-economic status.

The third level of research findings comes from studies that did control for specific intervening variables and systematically manipulated the label variable (Gottleib, 1974, 1975). The general procedures used involved the presentation of written description or videotapes of children identified as normal or handicapped depending on the experimental condition. After the presentation of the stimulus materials, subjects responded to semantic differential questionnaires on the observed children. Significant differences in ratings were obtained between experimental groups in which the type of label was the only variable (Seitz & Geske, 1976). They reported

that when children were not identified as retarded or emotionally disturbed, they received higher, more positive ratings than when they were identified. This clearly indicates a labeling effect, but the relevance of this finding is limited in terms of its implications for successfully integrating handicapped children. This was particularly apparent in the studies that combined the use of a label with behavioral data (Gottleib, 1974, 1975) presented via videotapes. When subjects rated the children in the film or predicted their abilities in certain tasks, the behaviors they observed accounted for much more of the variance than the label itself. One, extremely interesting finding (Gottleib, 1974, 1975) was that the specific behavior viewed influenced the direction of the differences in the ratings. For example, children seen engaging in academically competent and socially acceptable behavior received higher ratings than children viewed as incompetent and poorly behaved, regardless of the lable assigned to them.

In conclusion, one can say that labeling effects do exist, although, their scope and relevance to improving special education services is extremely limited. What appear to be the most important findings from this research is that perceptions of handicapped children which may be affecting attitudes and social adjustment, are

more likely determined by observations of a child's behavioral repertoire and noting how it contrasts or coincides with the repertoires of normal peers. Looking as the influence of specific classes of behavior in more detail has direct implications for effective mainstreaming. Research in this area may indicate which skills and social behaviors should be taught to a handicapped child to increase the likelihood of improved peer relationships and a more positive social status.

CHAPTER II

RATIONALE FOR THE PRESENT STUDY

The present study was designed in an attempt to evaluate the effectiveness of mainstreaming in a preschool setting. Of primary interest was the impact of integration on social adaptability of the handicapped children. In studying social adaptability, classes of behavior such as social interactions, language skills and play behaviors are relevant. Questions regarding which behaviors change as a result of integration and the types of changes that occur (e.g., increases or decreases in frequency) need to be considered. Of equal importance is the effect of mainstreaming on the nonhandicapped peers. As Bricker (1978) writes "if integration leads to a lack of productivity by the normal or above average youngster, the resources are not being effectively used" (1978, p. 15).

A variety of research methodologies are available for studying aspects of the mainstreaming issue. They include naturalistic or ethological observation procedures, direct interventions and indirect techniques.

Naturalistic Observation

Naturalistic observation methods are emphasized as they form the basis of many other methodologies and are particularly useful in studying classroom behavior. Repeated observations of handicapped and nonhandicapped students in an integrated setting provide a background against which observed change can be compared. As the composition of each classroom will be different, a description of the types and frequencies of behaviors that occur in that environment is essential for later evaluation purposes. In addition, a descriptive account of interactive behaviors occurring between the handicapped and nonhandicapped students is useful for designing intervention procedures. Before goals and objectives can be written, one needs to inventory the classes of behavior that are supported naturally by the environment.

Several descriptive studies have been conducted in integrated classrooms. Ray (Note 4), for example, found that nonretarded children interacted significantly more often with all students, handicapped and nonhandicapped alike. Porter, Ramsey, Trembley, Iaccobo, and Crawley (1978) observed that nonretarded children maintained the closest distance to nonretarded peers while retarded children were found to maintain the furthest distances. Frequencies of

social interactions also were recorded revealing that the nonretarded child showed a greater number of initiations with nonretarded than retarded peers. Similar findings were reported by Snyder, Apolloni and Cooke (1977) who summarized their data on the effect of an integrated preschool setting on social interactions in the following statement:

> studies...have indicated that integrated settings do not necessarily result in increased cross group imitation and social interaction between the handicapped and nonhandicapped children. Apparently, teaching procedures designed to foster these effects are needed if retarded and other handicapped children are to benefit optimally from integrated preschool programming. (1977, p. 263)

Naturalistic observation, then, is an important tool in describing ongoing behavior. It also can be used as an evaluation tool by comparing pre- and post-observations. For example, data could be collected on interaction patterns in the beginning of the school year and compared to measures taken at the end of the term. Specific changes in patterns of interactions of the handicapped and nonhandicapped students can then be assessed. One observational study (Novak,

Kearney, & Olley, Note 2) conducted in the preschool setting sampled play behaviors of special and nonspecial needs children during free-play periods. Differences were noted in types of play behavior and frequency of initiated contact between the two groups.

Reliability Assessment. A major issue in naturalistic observational research is the reliability of data collection techniques. Recently, the issue of observer reliability assessment has been a topic of discussion in professional journals. The tone of much of the discussion has been quite critical of percentage agreement methods so widely used in applied research. A series of articles in the Journal of Applied Behavior Analysis (Spring, 1977) illustrates the point. Yelton, Wildner, & Erickson (1977) noted three major weaknesses of percentage agreement: (1) it is affected by the frequency of behavior, (2) whether or not agreement is calculated on occurrences or nonoccurrences of the target behavior affects the coefficient, and (3) it does not take into account chance agreement between observers. Kratochwill and Wetzel (1977) added that an agreement coefficient may be (1) highly insensitive to the operational definitions, (2) may misrepresent observer competence, and (3) may not assess necessarily the "believability" of the experimental effect.

Judging from this sample of problems, it is apparent

that researchers must be careful when selecting a reliability assessment procedure. In spite of the criticisms levelled against percentage agreement, however, it is still a viable method depending on how and under what circumstances it is employed. This became quite obvious after examining the alternative procedures suggested to solve some of the problems. In general, the alternatives involve consideration of chance agreement in the calculation of reliability indices. Yelton et al. (1977), for example, present a probability-based formula that is sensitive to the number of observation intervals in the recording system. Its major drawback, though, is that it requires the use of a computer program or cumbersome calculations thus limiting its usefulness for spot reliability checks in the field. Hopkins and Hermann (1977) suggest that researchers report two separate series of reliability measures: a traditional group of observer agreement coefficients based on occurrences, non-occurrences, and overall measure, and a second set indicating the effects of chance agreement. Hartmann (1977) presents a case for using probability based statistics such as Kappa and Phi in place of percentage agreement. He suggests that these statistics may be more appropriate as they will reflect better the degree of reliability of the data.

In lieu of a probability-based formula, several

authors also recommend the use of effective percentage agreement when percentage calculations are employed (Hartmann, 1977; Hopkins & Hermann, 1977). The rationale for this suggestion is that effective percentage agreement provides a more conservative estimate of observer agreement as it reduces the possibility of spuriously inflated scores by eliminating either non-occurrence or occurrence intervals from the calculations (depending on the method used). For example, when recording a low frequency behavior the only intervals of concern are those in which the behavior occurs. If the non-occurrence intervals are counted as agreements, the agreement index will overestimate competence of the observers.

The implication of the collection of papers on reliability is that using a procedure that accounts for chance levels of agreement will provide for a better assessment of observer skill. In appreciation of the difficulties involved in collecting valid, believable, and reliable data the remarks by the various authors seem to address the problem and offer solutions. However, closer examination of the alternatives raises other issues which may be equally problematic. Of primary concern is the method used to estimate the chance agreement between observers. The formulae are based solely on the observed frequencies of the target behavior. If the behavior occurs at a very

high rate, then chance agreement is also going to be high. The measures of chance agreement will fluctuate across observations as the rate of behavior changes. Without a "standard" rate of behavior against which one can compare the observed rate it does not seem that knowing the probability of chance agreement can help a researcher collect more reliable data. This is a particular problem in settings where the rate of behavior fluctuates widely from one observation session to another. Baer (1977) in his rejoinder to Hartmann's (1977) paper makes this point and notes that behavior in the natural environment is not expected to be homogeneous. Instead, he suggests that what should be assessed is homogeneity across observers.

> Any number of observers equipped with the same definition of aggression should be able to look at the same event and say Yes or No homogeneously, if looking with that definition is a reliable process. That is where reliability is desired, so that is where homogeneity is meaningful. Hence, the homely measures of observer agreement so widely used in the field are exactly relevant to the problem (properly segregated for occurrence and nonoccurrence, of course). (p. 119)

For this particular study, effective percentage agreement based on occurrences was the procedure of choice. It enabled one to assess homogeneity across observers as well as provide a rapid feedback method to observers during training.

What percentage agreement cannot provide is information of the validity of the observational instrument. That is, even if two observers consistently agree does not mean necessarily that the same instrument can be used reliably in another setting by other observers. However, neither can probability-based formulae. As Hartmann (1977) notes, Kappa, Phi, percentage agreement and related statistics "completely confound random and systematic error". Therefore, the shortcomings of the percentage agreement method of assessing reliability must be counteracted by careful planning on the part of the researcher.

Naturalistic observational procedures can be used in conjunction with direct intervention procedures for the purposes of evaluating a program's impact. Collecting baseline performance data on handicapped and nonhandicapped students serves the same purpose as an observational study. It describes the current, naturally occurring patterns of behavior before a treatment program or intervention is implemented. Continuing to monitor behaviors after intervention enables an evaluation of the intervention's impact from an ecological perspective (Williams, 1974). Often a behavioral analysis report focuses data collection on specific behaviors targetted for change. In the case of mainstreaming, this might be the academic progress of the special needs student as most interventions or program

objectives are written for handicapped students. However, the intervention may have an impact on the nonhandicapped peers as well. Monitoring their behavior on the same variables during baseline and intervention would not only enable an ecological analysis but would permit identification of "spin-off" effects of the interventions.

A final note on the topic of naturalistic observation is its value in assessing social adaptability in a classroom setting. Observation of social interactions is a direct measure whereas other frequently used methods such as teacher opinions and peer ratings are indirect estimates. Gottleib (1978) points this out in the following statement.

> Since social adaptation is contingent on an individual's interactions with others it appears reasonable to suggest that measures of the adequacy of a retarded child's social adaptation should be based on the extent to which he demonstrates proficiency in interpersonal dealings. The only way to determine the adequacy of these interpersonal competencies is to observe them. (1978, p. 302)

Similarly, Porter et al. (1978) comment that in assessing social preferences, behavior recorded in the natural setting is more likely "a valid reflection of social preferences" than more commonly used sociometric questionnaires (1978, p. 324).

Direct Intervention Procedure.

A functional analysis of direct intervention forms a second category of research methodologies that are frequently used to assess skill achievement in both academic and social areas. In the academic sphere direct interventions for special needs students seem to fall into four categories: (1) special tutoring in a resource room, (2) individual tutoring by other professionals or paraprofessionals within the classroom, (3) peer tutoring, and (4) reciprocal peer tutoring. In each case, specific behaviors targetted for improvement are trained. Whether resource room personnel or itinerant teachers are available depends on the resources of the school system. Where resources are limited, peer tutoring procedures have been found to be extremely valuable. In addition to the cost-effectiveness of using students as trainers, peer tutoring and reciprocal peer programs appear to complement the goals of mainstreaming.

Peer tutoring has been used effectively to teach a wide variety of skills to children by all age groups (Feldman, Devin-Sheehan, & Allen, 1976). Tutoring has been shown to have positive social and academic effects on both the tutors and tutees. Children as young as preschool age have worked as tutors for their less able

peers (Cash & Evans, 1975; Long & Madsen, 1975). In addition, recent research has employed children with behavioral deficits or excesses as change agents. Niedermeyer and Ellis (1971) demonstrated the use of children with behavior problems as reading tutors. While the peer tutoring procedure has been shown to be effective, it is structured so that on child dominates the interaction by controlling stimulus materials and reinforcers. The reciprocal peer tutoring program on the other hand, involves equal participation of two children equivalently qualified with respect to the skills of concern. A recent report by Tinjaca and Goetz (Note 6) demonstrated the use of reciprocal tutoring programs to teach preschoolers a foreign language. In this particular study, one child was native Spanish speaking and the other spoke English. The authors pointed out that in addition to training the foreign language vocabulary the procedure was a cost-effective technique. It utilized students, the most abundant resource in the school, as trainers.

Peer tutoring has been found to have advantages beyond the academic achievements. Various authors have reported that children seem to find peer tutoring more enjoyable, have increased motivation, and improved self-concept and attitudes towards school (Yamamoto & Kentschy,

Note 7; Mohan, 1971). Peer tutoring also appears to benefit the tutor. Studies at the college level using the Keller Plan or Personalized System of Instruction (Keller & Sherman, 1974) have demonstrated that students who proctor other students' quizzes achieve higher scores on review test items than students who do not serve as proctors (Johnson & Sulzer-Azaroff, 1975). Another illustration is a study in which elementary school children with a two-year delay in reading achievement served as spelling tutors (Dineen, Clark, & Risley, 1977). The procedures involved three spelling lists. Each subject taught one spelling list, was tutored on another while the third list was not taught in any special way. A single subject design was used to compare individual subject pre-test scores to post-tutoring performances on the three lists. The results indicated that significant gains were made on the words they tutored and were taught.

The benefits to tutors and tutees in academic and effective realms, render peer tutoring an appealing procedure. Also, it can be cost-effective, making use of an abundant resource. Current research reports suggest that cross-age tutoring may be somewhat more effective than using tutors and tutees of the same age. Linton (1973) demonstrated that a two-grade differential in high school peer tutoring programs were more successful in changing

tutee behavior than same-age tutors. As mentioned, the peer tutoring so structured in a way so that one student, the tutor, is a "status" child and authority figure. Such a relationship may be more effective when the tutor is older. The reciprocal peer tutoring program, on the other hand, utilizes same age children who are equal participants in the interaction. As neither child is in a position of control over the other, it is likely that it would have a different effect on social interactions. Reciprocal peer tutoring also is logistically easier to manage for the classroom teacher. Cross-age tutoring would involve coordination of class schedules with other teachers whereas reciprocal peer tutoring can be arranged in individual classrooms. Hartup (1976) writes that same-age peer contacts may have more adaptational significance for children in complex societies. However, he also suggests that cross-age interactions may be beneficial for the handicapped child.

> The significance of same-age contacts is thought to derive from the more optimal balance of positive and negative feedback that occurs when children interact with age-mates than occurs during interaction of children of differing developmental levels. Nevertheless, deliberately designed cross-age interactions may facilitate specialization for children who have encountered certain kinds of developmental difficulties. (1974, p. 55)

Social skills training also has been accomplished using direct intervention procedures. Different methods include: (1) the reinforcement (or punishment) of a target behavior as it occurs in the natural environment, (2) training in general in a laboratory setting, and (3) intervention procedures designed to promote social skills improvements as "spin offs". The first method has been used frequently in classroom settings. Teacher attention or verbal praise contingently applied has effected changes in a variety of behaviors including sharing, following class rules, social isolation, etc. (Buell, Stoddard, Harris, & Baer, 1968; Rogers-Warren & Baer, 1976).

Specific social skills training programs are available for improving behaviors necessary for social interactions. A number of studies have been conducted with handicapped or behaviorally disordered children in an attempt to improve social interactions with nonhandicapped peers. Cooke and Apolloni (1976), for example, taught children to increase rates of smiling, sharing and positive verbal comments during social interactions. The results showed that in a free-play setting, trained children emitted the target behaviors more often than they had during baseline. In addition, untrained children in the same setting showed similar increases in these behaviors, presumably as a result of generalization, imitation, or a

combination of these and perhaps other factors.

Improvements in social adaptability often has been reported as a positive side effect of other intervention procedures. As mentioned, research on peer tutoring frequently has reported positive changes in the social interactions among tutors and tutees. Other procedures, such as the use of group contingencies have had positive effects on social behavior in the classroom. In a study by Frankowsky and Sulzer-Azaroff (Note 1), group contingencies generated positive verbal and non-verbal social behaviors. Contingency packages such as a modified version of the good behavior game used by teachers to handle disruptive students have resulted in positive social outcomes not specifically programmed, including student enjoyment of the game.

Reciprocal peer tutoring also can be viewed as a vehicle for developing social interaction skills. As mentioned, it is structured so that children participate at the same level. This is particularly relevant when considering how social skills develop. Patterson and Reid (1970) along with Strain and Shores (1977) and others propose that reciprocity is an important component in the development of social interaction skills. Reciprocity refers to a dyadic interaction in which the participants

reinforce each other at an equitable rate (Patterson & Reid, 1970). In the case of the handicapped child lacking adequate social interactions with normal peers, it is possible that specific behavioral deficits make the attainment of reciprocity highly unlikely. The reciprocal peer procedure may be useful as it provides an opportunity for the behaviorally deficient child to practice social behaviors important for the development of social interactions with normal peers.

Indirect Procedures.

In addition to direct intervention procedures, indirect techniques have been used frequently to assess social adaptability of handicapped children. Many of the efficacy studies cited earlier (e.g. Baldwin, 1958; Blatt, 1958; Shotel et al., 1972; Payne & Murray, 1974), represent examples of the use of indirect assessment methods. Data collection techniques included social status questionnaires, teacher and administration opinion polls, and peer rating scales. While the data have been interesting and informative, they only provide a general description of the status of the "handicapped" in the regular classroom. However, they do not offer insights into the specific conditions or factors that contribute to the social position of

the special children. Another problem is that they don't describe the types of behaviors or disabilities that were being rated by the normal peers, teachers or school principals. It would be useful for programming purposes if teachers of mainstreamed classes knew whether particular handicaps influenced the likelihood of achieving effective social integration more than others. If the impact of the disabilities on social acceptance were known, then plans could be made to attenuate any potential negative impace.

Using indirect procedures in conjunction with direct intervention would appear to be a useful collaboration. Direct procedures could determine the factors contributing to social status and indirect methods could be used to validate the choice of behavioral measures.

Present Study.

The focus of the present study was on the impact of mainstreaming on the social behaviors of handicapped children in a preschool setting. The term mainstreaming does not refer to one specific procedure. Rather, it is used as a general term and describes a variety of integration conditions found in the public schools. Integrated classes may vary according to size, ratio of handicapped

to nonhandicapped students, range of handicapping conditions, etc. Because of this diversity, the effects of mainstreaming must be examined on an individual classroom basis. Perhaps at some point in the future, the results of the individual evaluations can be summarized to provide a more global analysis of mainstreaming's impact. For now, the most productive approach would appear to be an analysis of the factors in individual classrooms relevant to accomplishing effective integration.

In choosing a method of evaluation at the single class level, one can focus on the children as a group, on all the children individually, or on specific children sampled to represent different subgroups. In the present study, the latter method was selected. Data were collected on two special needs students out of a possible eight receiving special education services, and on five of seven non-special needs children. Selection procedures are detailed in the following chapter. It was assumed that these children were representative of both mainstreamed and normal preschoolers.

In order to assess the impact of integration, naturalistic observation methods were used along with a direct intervention. Data were recorded on several classes of social behavior including verbalizations, social interac-

tions, and play behavior. After a baseline had been obtained on these behaviors the intervention was introduced. A reciprocal peer interaction procedure was selected as the method of choice considering the goals of the preschool, the needs of the students, and the available resources. The intervention titled the reciprocal skill development game employed reciprocal interaction in the form of taking turns in an activity and focused on language skills. Both ingredients, the ability to communicate and reciprocity are elements of spontaneous social interactions (Strain & Shores, 1977).

Observations of the three classes of behavior continued during intervention periods for both target special needs students and nonhandicapped peers. Comparisons to baseline data permitted several types of evaluation: (1) a determination of the effectiveness of the intervention as a procedure for developing social interaction skills in special needs children, (2) an assessment of the usefulness of the methodology in evaluating the program's effect, and (3) an ecological analysis of the impace of the intervention on the interaction patterns in the classroom.

CHAPTER III

METHOD

Subjects.

Participants in the study were 15 preschool students from an integrated classroom serving both special needs and non-special needs children. Two special needs students served as target subjects whose behavior was tracked during all phases of the experiment.

Target Subject A: This student was a 4.3 year old female who was identified as a special needs student on the basis of her social/emotional behavior. She exhibited low frequencies of social interactions or verbalizations with peers and adults and would generally be categorized as withdrawn. On standardized measures of cognitive ability, she scores within the average range of intelligence.

Target Subject B: This student was a 6.4 year old male with Down's Syndrome. His social behavior was characterized by low frequencies of vocal expressive language and high frequencies of gestural communication. He interacted regularly with his peers and often initiated the interactions. On measures of cognitive ability, Subject B scored well below the mean for his age, indicating severe developmental delay.

Special services for both subjects consisted of participation in the mainstreamed preschool program on a daily basis and speech therapy. Speech services were delivered individually several times a week by the speech pathologist.

Subject selection was based on several criteria. These students satisfied the following conditions: were full-time students, had been enrolled in the program since the beginning of the school year, were receiving individual services, and whose special needs were in the area of social behavior.

Non-special Needs Students: Other participants in the study were the non-special needs peers of the target subjects. Of the fifteen children who were enrolled in the program, eight were not receiving special services. A breakdown of students by age, sex, educational placement, and level of attendance can be found in Table 1.

Informed Consent.

Written permission was obtained from all parents of all students participating in the study. A cover letter describing the project was mailed to parents along with a consent form which was to be returned to the experimenter. Of the 15 students in the class, the parents of only one

Table 1: Distribution of students in the preschool program by age, sex, student status, and attendance.

Student	Sex	Age	Non-Special Needs	Special Needs	Full Time	Part Time
1	M	3.1 yrs		X		X
2	M	3.5 yrs	X		X	
3	M	3.6 yrs		X	X	
4	M	4.2 yrs		X	X	
5	M	4.3 yrs	X		X	
6	M	4.8 yrs	X		X	
7	M	4.9 yrs		X		X
8	M	6.4 yrs		X	X	
9	M	7.9 yrs		X		X
10	F	3.2 yrs		X		X
11	F	3.3 yrs	X		X	
12	F	4.0 yrs	X		X	
13	F	4.3 yrs		X	X	
14	F	4.4 yrs	X		X	
15	F	4.6 yrs	X		X	
Summary	M=9 F=6	Median Age= 4.3 yrs	7	8	11	4

boy refused to give permission for his participation.

Setting.

The Early Education Center, an integrated preschool program housed in a local elementary school was the setting for the study. The program was designed to accomodate special needs and non-special needs students in compliance with state and federal legislation (Massachusetts Chapter 766 and Public Law 94-142). The Early Education Center is supported by a federal grant and is staffed by a teacher, an assistant teacher, a speech pathologist, and a preschool coordinator. In addition, several student interms from the special education and communication disorders departments of the University of Massachusetts, assisted in the operation of the program.

Approximately half of the students enrolled in the program were designated as special needs. Identification procedures for locating special needs children were available within the community. This included an area-wide preschool screening program conducted by a team of speech pathology and psychology graduate students, a nurse, a certified speech pathologist, and a special educator. Any child suspected of having special needs was then given an in-depth diagnostic evaluation. The evaluation process

involved the cooperative efforts of the following persons: the parents, a psychologist, speech pathologist, physician or nurse, and a person certified to conduct a developmental and family history. For those diagnosed as having special needs, special goals and objectives were determined for the child's educational plan.

Parents of non-special needs children interested in enrolling their children in the program were required to complete an application and be interviewed by the program coordinator. Final selection of students to fill the available positions occurred by means of a lottery.

The Early Education Center consists of a large classroom divided into a number of activity centers such as: library corner, an office area, a climbing structure, housekeeping corner, block corner, an area for table top activities, a painting area, an open space for large group activities, and sand, water, and mud tables. A diagram of the classroom is presented in Figure 1. Available in each of the activity centers is a varied assortment of toys, games, and materials for the students' use.

The program operates on a regular daily schedule beginning at 9:00 a.m. and ending at 11:45 a.m. The approximate times alloted for the scheduled activities are as follows:

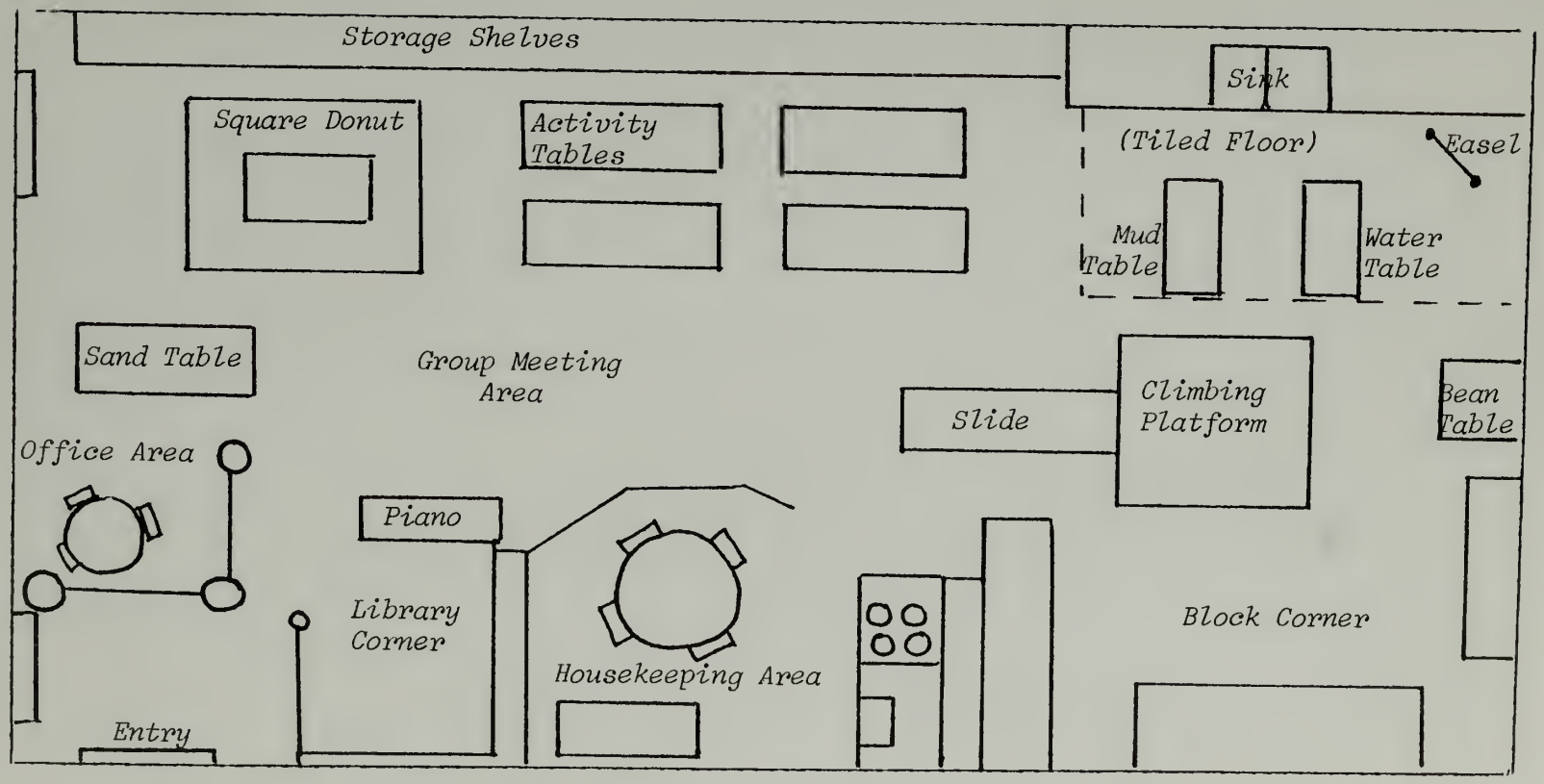


Fig. 1: Schematic drawing of the preschool classroom.

9:00 - 9:15 Morning Circle

9:15 - 9:45 Teacher Directed Activity

9:45 - 10:30 Free Play Period

10:30 - 10:45 Snack

10:45 - 11:00 Story Time

11:00 - 11:30 Recess

11:30 - 11:45 Singing/Music Period

The present study was conducted during the teacher directed activity and free-play periods. Children worked in pairs with the experimenter for 15 minutes at the start of the teacher directed activity period. These brief intervention sessions were conducted in the carpeted hallway adjacent to the classroom. (This is a location frequently used by teachers for small group lessons.) Observations of the children's behavior were conducted primarily during the free-play time, however, occasional observations were taken during the end of teacher-directed activities.

Project Approval.

Consent to conduct this research project was granted by both the preschool coordinator and the classroom teacher. The project was then approved by the Board of the Regional School District and the Human Subjects Committee at the University of Massachusetts.

Research Personnel.

The study was conducted by the experimenter, a doctoral candidate in Psychology and a trained research assistant. The assistant was a psychology major with previous experience working with handicapped populations. The research assistant's primary task was data collection. To prevent observer bias while recording, the assistant was not informed fully about the nature of the investigation. Involvement of the preschool personnel was limited. Only the program coordinator and classroom teacher were aware of the purpose of the study and their participation basically consisted of introducing the experimenter to the children and explaining that most students would have the opportunity to play a game with her.

Apparatus.

Intervention: Reciprocal Pre-academic Skill Development Game. Required materials for the game included the game board, a deck of picture cards, small toy animals to serve as game pieces, and a spinner. The game, modelled after Candy Land (Milton Bradley, Inc.) a popular commercial game for preschoolers, displayed a circus theme. Fifteen red, blue, green, and yellow colored spaces formed a winding path leading from the entrance to a circus "Big

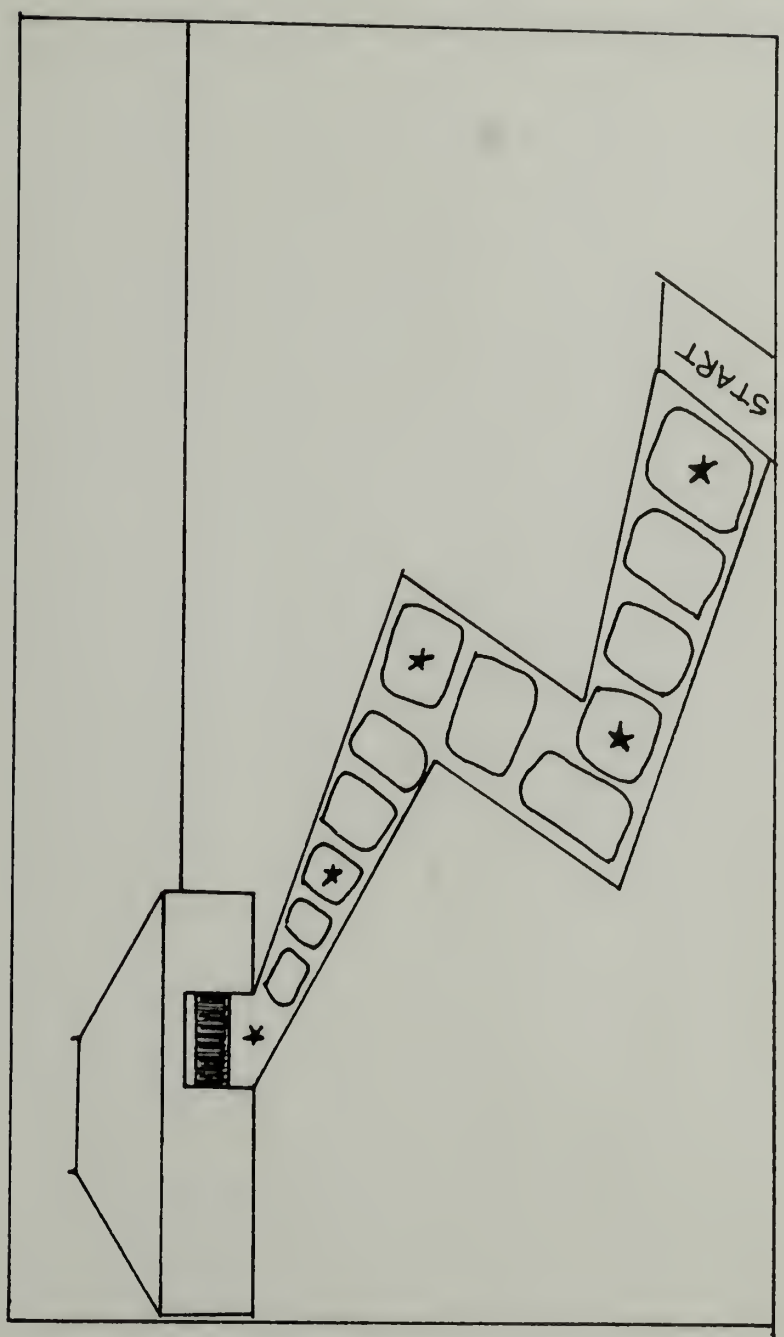


FIG 2: DIAGRAM OF GAME BOARD

Top". (See Fig. 2) Each of the four colors appeared on the spinner which was used to direct the movement of the player's game piece along the colored path. Opportunities to spin were earned by correctly describing a picture card drawn from the opponents hand. (The rules of the game will be described in detail in the Procedure Section.) The cards depicted various action scenes such as a man changing a flat tire, a boy sliding down a slide, or a girl pouring juice into a glass. Ten cards were used in all.

Reinforcers: During the game, children had the occasion to earn small tangible reinforcers in the form of stickers and stamps. Reinforcers were awarded when a player landed on a specially designated space. Of the 20 spaces on the board, 10 were marked with stars indicating that reinforcements were available. Whenever a child landed on a "starred" space, he or she could choose one sticker or stamp from an assortment of over 20 different pictures and shapes.

For each game played, participants received a personalized sheet of paper on which they could affix their stickers. Most children exclusively used their papers for displaying their earnings, however, several innovative children found more interesting places to show their prizes

such as their faces, clothing, shoes, and even stomachs.

In addition to the tangible reinforcers, subjects were given verbal praise and pats from the experimenter. Children were also prompted to provide each other with positive feedback throughout the intervention.

Data Collection: The equipment required for data collection included audio and video tape recorders, stop watches, and specially prepared data sheets. All the intervention sessions were recorded on audio tape. The video tapes, on the other hand, were used only as a means of documenting the behavior during baseline and following the implementation of the social interaction game, and as a method of checking observer reliability. The first video recording was taped during the first week of baseline and the second occurred four weeks after intervention procedures had begun. The stop watches and data sheets were used by observers to record three classes of behavior during the free play period.

Observation System.

A partial interval scoring system was designed to record three classes of behavior: vocalizations, social interactions, and play behavior. Within each class, several subclasses of behavior were monitored. Fifteen second in-

tervals were used and observation periods were four minutes long. The specific subclasses and the coding systems used for each category are defined below. Many of the behavioral definitions were adapted from earlier research in the field (Porter et al., 1978; Novak et al., Note 2).

Vocalizations. In this category, the vocal verbal behavior of the target subject was recorded. This included initiated statements as well as responses to the statements of others.

Behavior Code: Vocalizations

(I) Initiation: Any utterance directed to another individual as evidenced by one or more of the following: head turned towards addressee, eye contact, use of addressee's name, or content of the statement. Only coded as (I) if subject speaks after at least 5 seconds of no talking to a peer or adult. If, however, subject speaks to a second person within the 5 second limit, the statement may be scored as (I).

(R) Response: Any utterance made in response to a direct question or statement from a

peer or an adult. To score as (R), the reply must occur within 5 seconds of the verbal antecedent.

(RC) Response/Conversation: Vocalizations should be scored (RC) if the subject's response is followed by a continuation of initiations and responses between the subject and listener(s). Subject must make at least two topically related responses within a 15-second interval to qualify as (RC). If two responses are made to two separate verbal prompts in the same interval do not score as (RC) but as (R), (R).

( ) Initials: For each utterance coded as (I), (R), or (RC), add a subscript with the initials of the participant in the verbal interchange. For example, (I) followed by (AB) would indicate a person with the initials A.B. was addressed by the subject.

Behavior Categories: Vocatlizations

To Adult: Any utterance directed to the teacher, a staff member, or other adult present in the classroom (e.g., parents, re-Research staff, visitors). Utterances should include all statements, questions,

answers, exclamations, etc. Does not include laughing, crying, screaming, or other noises. Be sure to add initials of the adult involved in the interaction.

To Peer: Any utterance directed to another student in the class or student from a neighboring classroom. Includes statements, questions, answers, etc. Be sure to include subscript indicating who was involved in the interaction.

Ambiguous: Any vocalization that is non-directive or unintelligible. This includes instances of subject talking to him or herself, making noises, or producing sound effects while playing.

Other: Any other vocalization not defined above such as laughing, crying, or screaming. Be sure to specify the specific behavior when using this category.

Social Interactions. This class of behavior consisted of the physical and social interchanges that occurred between the target child and another peer or adult. The specific behaviors appearing in this category were derived from related research (Porter et al., 1978; Kearney, 1979) and from preliminary observations in the preschool setting. Represented in the list were both positive and negative

affective behaviors and social interactions resulting from both subject and peer initiation.

Behavior Code: Social Interactions

(A) Adult: Use to indicate that an interaction occurred between the subject and an adult.

(P) Peer: Use to indicate that an interaction occurred with a peer.

( ) Initials: The initials of the peer and/or adult with whom the interaction occurred are to be included as a subscript to the letter code above.

Behavior Categories: Social Interactions

Give: An object held in child's hand is held out for another person to grasp and is then released. Includes objects that are placed in a lap or directly in front of another person.

Offer: An object held in child's hand is held out towards another person but is not released.

Affection: Includes any behavior which brings individuals into physical contact in a non-aggressive, friendly manner such as a hug, kiss, pat, holding hands, linking arms, arms around one another, etc.

Aggression: Includes any behavior which brings

individuals into physical contact in a rough and unfriendly manner such as hitting, slapping, shoving, poking, punching, yanking, etc.

Receive: Child accepts an object given or offered by another person. Includes items that are placed in subject's lap, directly in front of subject, and objects taken directly into subject's hands.

Take-Tug-Grab: Taking an object from someone's hands which has not been offered. Grab is a faster version of take and often includes resistence. Tug involves holding onto and pulling an object which someone else is holding and does not immediately release.

Other: Other social interactions not defined above. When using this category, be sure to record the specific behavior observed and the participants.

Play Behavior. This category was used to document the type of play activity the subject was engaged in during observation periods. Within a free play session the subject's behavior could be described as one of the following: solitary play, parallel play, cooperative play, passive watching, with adult, or other. If observations were

conducted during the teacher directed activity period, play behavior was coded as such.

Behavior Code: Play Behavior

( ) A check mark was used to indicate the type of play behavior the subject was engaged in during each 15 second interval.

Behavior Categories: Play Behavior

Solitary Play: Child is actively engaged in independent play and is not within 3' of another child.

Parallel Play: Children playing in close proximity to each other (within 3') yet are playing independently. The subject does not attempt to influence or direct the play of others. There is very little interaction among the children.

Cooperative Play: Subject is playing with at least one other child. Interactions must occur and may include conversation, joint activity (e.g., pushing a child in the wheelbarrow, rocking a child in the hammock), borrowing and sharing toys, following and chasing one another around the room, or other organized play without an adult present.

Passive Watch: Child observes the behavior of

others for at least 5 seconds and is not simultaneously engaged in play. Includes instances where child disengages from an activity to watch others. Only score if subject's observation lasts a minimum of 5 seconds.

With Adult: If an adult (teacher, staff member, parent) is present in an activity area with the target child and/or a group of children, and is participating, directing, or is otherwise involved in the activity, score as "with adult". If adult is merely watching from a distance and makes no attempt to intervene, physically or verbally, score the type of play behavior that is observed.

Other: Child is in a state of transition, such as changing play activities, washing hands, wandering around the room, etc. The category should also be used in the event the child momentarily leaves the room or does something other than play. Be sure to specify the particular behavior that has occurred.

Teacher Activity: To be used if observations are conducted during the teacher activity period which precedes free play. Should be used to

indicate that one teacher is working on an activity (generally at the activity tables) with a small group of students. If no teacher is present at the table or area, then score as if behavior was occurring in the free play period.

Data sheets with the three classes of behavior and all of the subclasses were prepared. Each sheet represented one four-minute observation period with each minute divided into four 15-second intervals. A sample sheet is presented on page 106. Note that for each observation, the time, setting, subject, and observer are recorded. Also space is provided for notes and other pertinent comments.

Observer Training.

Observer training was conducted in three stages, each one requiring more complex behavior from the observer. The purpose of utilizing a graduated training program was to shape increasingly sophisticated observation skills. The format employed was adapted from the observational teaching procedures developed by Reese (Note 5) and Novak (personal communication) both of whom have developed programs to train naturalistic observational recording.

The first level of training involved the introduction

Subject: Date: Time: Observer:

Setting:

Vocalizations Initiated R = Response R/C = Response Conversation

	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60
To Adult																
To Peer																
Ambiguous																
Other																

Social Interactions A = Adult P = Peer

	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60
Give																
Offer																
Affectionate																
Aggressive																
Receive																
Take-Tug Grab																
Other specify)																

Play Behavior √ = indicate type of behavior S is engaged in

	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60
Solitary																
Parallel																
Cooperative																
Passive Watch																
Teacher Direc.																
With Adult																
Other Specify)																

Notes:

Sample Data Sheet

of the behavioral categories and coding systems. Operational definitions were studied by the assistant and discussed. Hypothetical recording problems were presented which the observer had to score. Following this introduction, on-site observations were conducted. At this stage of training, though, the goal of the observation was merely to recognize instances of the different target behaviors as they occurred. Any disagreements or difficulties encountered were resolved through discussion and reference to the operational definitions.

The second stage was devoted to practice recording with the data sheets and stop watches from video-taped samples of classroom behavior. Tapes included the free play behavior of both target subjects. Two observers would record simultaneously a three minute sample of the tape. The advantage of this procedure was that it enabled the observers to re-examine the behaviors in the event of scoring disagreements and thus facilitated clarification of the behavioral code. Observer agreement indices based on the occurrence of the target behaviors were calculated for all practice recordings. Hartmann (1977) termed this procedure effective percent agreement which is represented by the following formula.

$$\text{Effective Percent Agreement for Occurrences} = \frac{\#\text{ Agreements}}{\#\text{ Agreements} + \#\text{ Disagreements}}$$

Video tape practice remained in effect until observers consistently reached an agreement index of 85%.

During phase three, simultaneous recordings were conducted in the preschool classroom. Immediately following each observation, the observer and trainer reviewed the data sheets and computed an agreement coefficient checking each scoring category. Decisions made regarding the scoring of a particular behavior were noted on the behavior code sheet for future reference. These decisions could then be reviewed prior to the start of daily observations. Approximately three weeks were spent in recording practice at which time agreement levels averaged 85%.

Schedule of Observations.

Observations were conducted during the 45-minute free-play period, three times a week. On each day, target children were observed as many times as possible on a randomized schedule. Attempts were made to conduct observations on an alternate basis, that is observing child A, then child B, and back to child A, etc. However, this sequence was not always possible as absences, special activities, and other unanticiapted events often interfered. For example, both target subjects received their speech services several times a week during free-play which would take them away

from their play activities. On other occasions, the teacher directed activity period that immediately precedes free-play ran longer than usual thus shortening observation time. Generally, it was possible to obtain two or three observations of each target subject each day.

Periodically, the play behavior of the non-special needs students was recorded. Observations were made randomly, when either of the target children were not available. The purpose of these observations was to provide a basis for comparison of the play behavior of the special nedds and non-special needs students.

Reliability.

Interobserver reliability was assessed using an effective percentage agreement formula based on occurrences of the target behavior (Hartmann, 1977). This procedure was selected after consideration of its advantages and limitations compared with those of available alternative procedures. In the present study, several precautions were taken to maintain the validity of the observation instrument.

Observer Bias. During all phases of the investigation, the observer was unaware of the schedule of intervention and was not fully informed of the purpose of the

experiment. Daily observations did not commence until the children had been returned to the classroom following intervention.

Observer Drift. To prevent observer drift, the general shift in criteria for judging the target behavior (Hersen & Barlow, 1976), two procedures were employed. One involved the frequent review of all operational definitions and all ambiguous recording situations encountered during previous observations. The second consisted of frequent unannounced reliability checks. Research (Hersen & Barlow, 1976) has demonstrated that this procedure results in higher accuracy levels.

Reactivity. Both subject and observer reactivity are potential problems when collecting data in a natural setting (JABA, 1977; Hersen & Barlow, 1976). In the present study, subject reactivity was not a problem in spite of the fact that observers were clearly visible and moved around the classroom. This fortunate circumstance seemed to be a result of the classroom's "open door" policy towards visitors. Since the beginning of the year, students have been accustomed to frequent adult visitors, including parents, student teachers, and other interested persons. Secondly, the experimenter had been a regular observer throughout the first semester, thus, a familiar figure to the students.

And, finally, efforts were made by the observers to be as inconspicuous as possible while recording and to prevent the subjects from learning who was the object of the observation. To assist in this measure, feedback on obtrusiveness was provided by the classroom teacher.

To minimize reactivity of observers and to insure independent recordings, observers were separated by a minimum of approximately eight feet when observing a child.

The mean percentage of observer agreement was 84%. For subject A, 33% of all observations were checked for reliability, whereas, 58% of subject B's observations were checked. Approximately 10% of observations of the non-special needs children were checked. The mean agreement percentages for each subject were 86% and 82% respectively. The range of coefficients was from 40% to 100% agreement. Observer agreement was calculated on 10% of the data on non-special needs subjects. Agreement percentages ranged from 52% to 93% with a mean of 81%. Reliability checks occurred during all phases of the experiment.

Experimental Design for Analyzing Effects of the Skill Development Game.

A multiple-baseline across subjects design was used to determine the effects of the skill development game.

This design permits sequential introduction of the intervention with different subjects and is particularly useful when treatment variables are irreversible or cannot be withdrawn due to practical or ethical considerations (Hersen & Barlow, 1976). In the present investigation, carryover from the treatment phase was anticipated thus eliminating the possibility of using the more rigorous reversal design. The multiple-baseline design, in this case, can be conceptualized as two A-B designs with different length baselines. The design is schematically presented in Fig. 3.

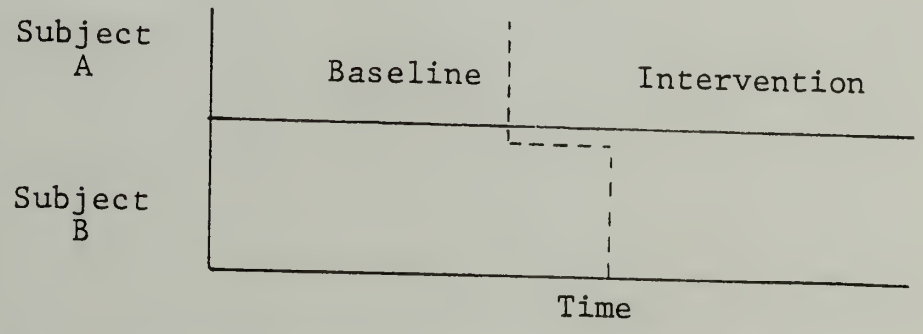


Fig. 3: Multiple-baseline design across two subjects.

Evidence of an experimental effect is present when changes in the baseline rates of behavior occur in conjunction with the introduction of the treatment variable.

Procedure.

Baseline. The baseline period for Subject A was three

weeks, during which time eleven days of observation were obtained. The baseline for target Subject B extended an additional two weeks in which 15 observation days were obtained. On any given day, observers alternately recorded the behavior of the two target children during the free-play period. On the average three observations per target subject were acquired on each day. Observations were conducted a minimum of three days a week. Occasionally, the number of observations was limited as a result of several uncontrollable variables. These included absences, late arrivals, over-extended teacher directed activity periods, and speech therapy sessions.

Intervention. Prior to the first day of intervention, the classroom teacher had announced that a number of children would be able to work with the experimenter and play a circus game. Then, on the first day and each succeeding day of intervention, the teacher would "select" two students to play the game during the announcement of the daily activities in morning circle. Children were actually predesignated by the experimenter. The children were then accompanied by the experimenter to the work area in the adjacent hallway.

Peer Selection. Of the seven non-special needs children in the program, six had permission to participate in

the study. One female student, however, did not wish to be involved, leaving five available participants (two male; three female). Each day, one of these five students was invited to play the game with target subject A or B. A schedule was devised so all children would play equally often and be evenly distributed throughout the intervention period. Occasional adjustments were required to accommodate student absenteeism and special classroom events.

Before beginning play, peers were pre-tested for color discrimination, identification of animals used as game pieces, and identification of several circus pictures such as the monkeys, clowns, balloons, and the tent. This procedure was included to insure basic entering behaviors necessary to play the game. All peers who served as subjects possessed these minimum skills.

Orientation. At the start of each intervention session, several minutes were reserved for orientation to the game and the setting. The experimenter chatted freely with the children, arranged materials, distributed cards, and in general, created an informal, relaxed atmosphere. This time also proved necessary to allow the children to get settled. They usually were very excited about coming out to play the game and would quiet down a bit before a

play began.

After players were settled, a brief explanation (or review) of the rules and sequence of events was provided. Essentially, children were reminded that (1) the object was to move their animals along the colored path to the circus tent, and (2) they would have to describe the pictures on the cards in order to get a turn to move their animals.

Playing the Game. Before starting play, each child had five picture cards and a toy animal. The animals were placed on the start space and one child was chosen to go first. At the completion of his or her turn (the following seven steps), the second player would go. Play alternated between players until all cards had been used or one child reached the circus tent.

(1) Player #1 selects a card from opponents hand

(2) Player #1 describes the scene depicted on the card

(3) Card is shown to Player #2 for approval

(4) If description is accurate and meets approval, Player #1 spins the spinner

(5) Player #1 moves his animal to the appropriate colored space

(6) If the space has a "star", the player selects a

sticker or other reinforcer from the reinforcement box held by the opponent

(7) Player #1's turn is complete and he holds his cards for Player #2. Steps 1 - 6 are repeated.

Special training sessions were not utulized in teaching the game. Instead, children were taught as they played. That is, each step of the play sequence was prompted by the experimenter for the first few turns. Thereafter, prompts were provided only when necessary to maintain continuity in game play. Subject B, in particular, required frequent prompting with statements such as "Who's turn is it?", or more direct comments like "Hold up your cards". All intervention sessions were tape recorded and averaged 15 minutes.

Schedule of Intervention Sessions. The original schedule allowed for eight exposures to the social interaction game delivered in two, four-session time blocks. Unfortunately, student absenced necessitated a revision which is shown in Table 2.

Table 2. Schedule of intervention sessions with target Subjects A and B during the eight weeks following baseline

	Weeks following Baseline							
	1	2	3	4	5	6	7	8
Subject A	1,2	3,4			5,6			
Subject B			1,2	3			4,5	6

As mentioned earlier, intervention sessions occurred during the teacher activity period. At the conclusion of the game, children returned to the classroom and resumed daily activities.

CHAPTER IV

RESULTS

The data on vocalizations, social interactions and play behavior collected during free-play periods constitute the major results of the study. Data were also obtained during the intervention sessions and will be presented first. Following the intervention data, the results are shown for each of the three major dependent variables by individual target subjects. After the individual summaries, group results are presented for the nonhandicapped students and comparisons are made.

Intervention

Before examining the impact of the skill development game on free-play behavior, subjects' participation in the game was evaluated. Participation required that players follow a sequence of behaviors on each turn including describing a picture card. How well the target children learned the behavior chain was assessed by the number of adult prompts required to maintain the sequence. Two types of prompts were recorded: (1) vocalization prompts - This type of prompt was a non-informational one provided

to elicit verbal descriptions of the cards. Examples included questions such as "What's on the card?", "What is the boy doing?" etc. (2) sequence prompts - Those were noninformational cues used to maintain the behavior chain of each turn. Examples of sequence prompts used are "What happens next?", "What do you do now?" or more direct statements such as "Who's turn is it?" Prompts were provided if a subject did not respond within seconds of selecting a card or completing a step in the game sequence. Prompts were delivered by the experimenter and recorded by audio tapes.

The results showed that sequence prompts were necessary for both target children in all intervention periods. Vocalization prompts were needed for Subject B but not for Subject A. These data are shown in Table 3 as percentages. The percentages indicate the mean ratio of prompted responses to non-prompted responses.

Table 3: Percentages of prompts used during intervention sessions for Target Subjects A and B.

Prompt Type	Subject A	Subject B
Vocalization	0%	78%
Sequence	42%	83%

These data represent the mean prompting rates for all intervention sessions. For Subject A, the mean level is representative of her performance during each session. No changes in trend were evident over time. For Subject B, some improvement was evident as vocalization prompts decreased over time. Spontaneous descriptions occurred during the final two intervention sessions accounting for two of the five trials in each game. This was an increase from no unprompted vocalizations during the initial intervention periods. For both subjects prompts successfully elicited verbal responses.

The following sections report the results obtained from free-play period observations. Vocalization data are shown first after which social interaction and play behavior data are presented.

Vocalizations

Subject A. The percentage of occurrence was calculated for the classes of behavior "initiation" and "response" and represents the number of intervals scored for a particular behavior divided by the total number of intervals observed. These data were collected during free-play periods. The results for Subject A are shown in Tables 4, 5, and 6, and Figures 4 and 5. Table 6 presents the ratio of

responses to the number of opportunities the subject had to respond. Opportunities consisted of initiations from peers or adults directed to the subject. Both the raw frequency and percentages are shown. Data were summarized across baseline and intervention phases due to the low frequency of occurrence of the target behaviors.

Table 4: Percentage of occurrence of Subject A's "initiations" during baseline and intervention.

	Initiations	
	Baseline	Intervention
To Adult	.29%	1.65%
To Peer	.89%	1.15%

Table 5: Percentage of Subject A's "responses during baseline and intervention

	Responses	
	Baseline	Intervention
To Adult	1.78%	1.49%
To Peer	0%	.66%

Table 6: Ratio of Subject A's responses to the number of opportunities to respond during baseline and intervention

	Responses/Opportunities to Respond	
	Baseline	Intervention
To Adult	6/13 = 46%	9/20 = 45%
To Peer	0/8 = 0%	4/8 = 50%

During baseline, initiations to peers occurred at a slightly higher rate than to adults. Responses, on the other hand, showed the opposite trend. Responses to adults were markedly higher than to peers. Following intervention, several changes in the rates of behavior were noted. Initiations towards peers increased as did initiations to adults. Responses to peers also showed an increase although, responses to adults decreased. The ratio of responses to opportunities indicates, however, that Subject A maintained approximately the same percentage of responses to adult initiations. The data on peer ratios supports the increase in responses to peers found in Table 5.

Within the category of "other" vocalizations, some interesting data were revealed for the specific behaviors of smiling and laughing. These data are shown in Fig. 6 and Table 7.

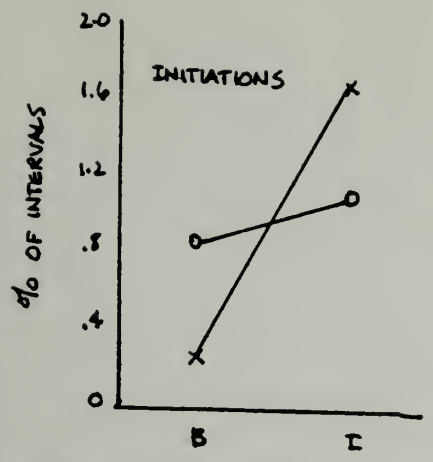


FIG 4: % OF INTERVALS OF SUBJECT A'S INITIATIONS TO ADULTS AND PEERS DURING BASELINE AND INTERVENTION

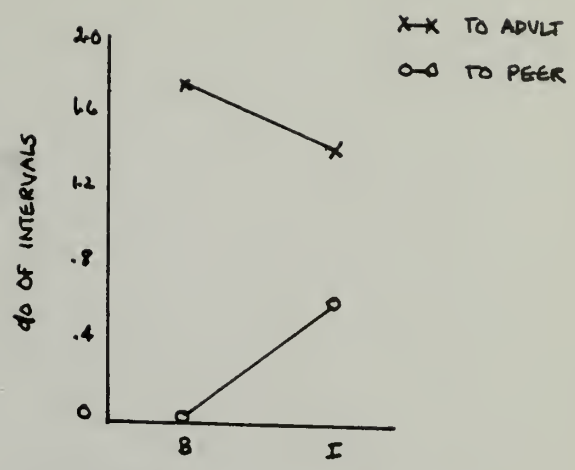


FIG 5: % OF INTERVALS OF SUBJECT A'S RESPONSES TO ADULTS AND PEERS DURING BASELINE AND INTERVENTION.

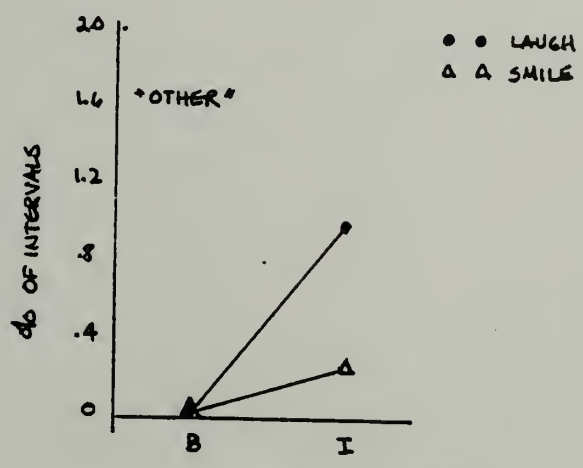


FIG. 6: % OF INTERVALS OF SUBJECT A'S "OTHER" VOCALIZATIONS (SMILE AND LAUGH) DURING BASELINE AND INTERVENTION.

Table 7: Percentage of Subject A's smiling and laughing behavior during baseline and intervention

	Baseline	Intervention
Smile	.0	.99% Percent
Laugh	.0	.23% Percent

Daily frequencies recorded for both classes of behavior can be found in Appendix III.

These positive affective vocal behaviors were only observed during the intervention phase of the investigation

Distribution of Vocalizations. Vocalizations recorded during free-play periods were coded with the initials of the participants in the interactions. The number of different students and adults involved in vocal interactions with Subject A are shown in Table 8.

Table 8: Number of individual peers and adults involved in vocal interactions with Subject A.

	Baseline		Intervention	
	Initiations	Responses	Initiations	Responses
To Adult	2	2	2	4
To Peer	2(1sp, 1nsp)*	0	4(1sp, 3nsp)	1(1nsp)

*sp = special needs student nsp = non special needs student

Content Analysis. The function of Subject A's vocalizations were examined according to Skinner's (1957) classification of verbal behavior. The analysis was restricted to vocal verbal behavior although it is recognized that "any movement capable of affecting another organism may be verbal" (Skinner, 1957, p. 14).

The major portion of Subject A's vocal verbal behavior during baseline and intervention consisted of what Skinner has termed "mands". A mand is a "verbal operant in which the response is reinforced by a characteristic consequence and is therefore under the functional control of relevant conditions of deprivation or aversive stimulation" (Skinner, 1957, p. 36). Mands can also be thought of as requests, commands, entreaty, and questions. This class of verbal operant is controlled by the listener unlike other verbal operants in Skinner's functional analysis which are under the control of the physical environment (e.g., the tact). Other classes of verbal behavior present in Subject A's repertoire included the echoic, nomination, and the intraverbal. Table 9 presents the frequency of the different verbal operants during baseline and intervention. In addition, examples of the operants are provided.

Table 9: Frequency and classification of Subject A's vocalizations.

Subject A's Vocal Verbal Behavior	Classification	Frequency B	Frequency I
"Look"	Mand	4	10
"Stop it"	Mand	1	2
"I want to do that"	Mand/Tact	4	2
Request for assistance	Mand	0	3
Request for materials	Mand/Tact	0	2
Naming Objects	Nomination	0	2
Repeating Words (e.g., yes, no, yes, no)	Echoic	0	2
"I did enough", and other responses to verbal stimuli	Intraverbal	1	2
	Total	10	25

It was apparent that the overall frequency of verbal behavior increased during intervention as did the range of verbal operants. The greatest change, however, occurred within the class of mands where the baseline total was 9 as compared with the intervention total of 19.

Subject B. The percentages of Subject B's vocalizations are shown in Figures 7 and 8 and Tables 10, 11, and 12.

Table 10: Percentages of occurrence of Subject B's "initiations" during baseline and intervention

	Initiations	
	Baseline	Intervention
To Adult	6.7%	3.7%
To Peer	4.1%	3.4%

Table 11: Percentages of Subject B's "responses" during baseline and intervention

	Responses	
	Baseline	Intervention
To Adult	5.0%	2.8%
To Peer	.7%	0%

Table 12: Ratio of Subject B's responses to the number of opportunities to respond during baseline and intervention

	Responses/Opportunities to Respond	
	Baseline	Intervention
To Adult	21/39 = 53.8%	13/20 = 65%
To Peer	3/8 = 37.5%	0/1 = 0%

In addition to initiations and responses, the "ambiguous" vocalizations for Subject B are presented in Figure 9 and Table 13. These data are included as they accounted for a large proportion of the total vocalizations.

Table 13: Percentage of "Ambiguous" verbal behavior during baseline and intervention for Subject B.

Ambiguous	
Baseline	Intervention
17.1%	23.1%

Both initiations and responses showed higher response rates during baseline. Although the frequency of responses to adults decreased during intervention, the ratio of responses to response opportunities increased. This indicates that Subject B responded to a higher proportion of adult initiations than he had during baseline. These behaviors also occurred more frequently towards adults than towards peers in both baseline and intervention phases. These data show an opposite trend than those for Subject A. However, the overall percentage of vocalizations towards both peers and adults was higher for Subject B. That is, his baseline levels of vocalizations were higher than Subject A's. Off-

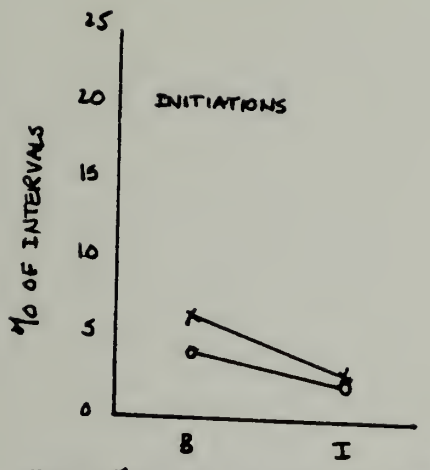


FIG 7: % OF INTERVALS OF SUBJECT B's INITIATIONS TO PEERS AND ADULTS DURING BASELINE AND INTERVENTION

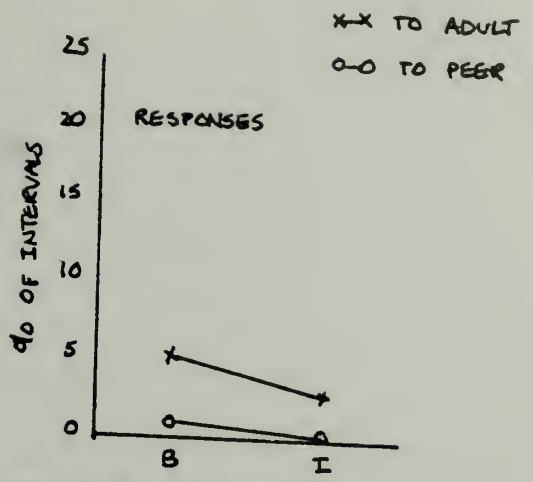


FIG 8: % OF INTERVALS OF SUBJECT B's RESPONSES DURING BASELINE AND INTERVENTION TO PEERS AND ADULTS

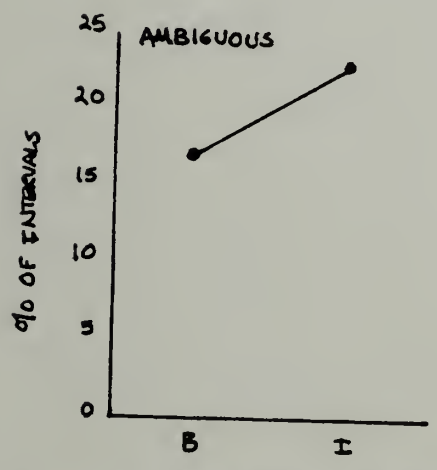


FIG. 9: % OF INTERVALS OF SUBJECT B's AMBIGUOUS VOCALIZATIONS DURING BASELINE AND INTERVENTION.

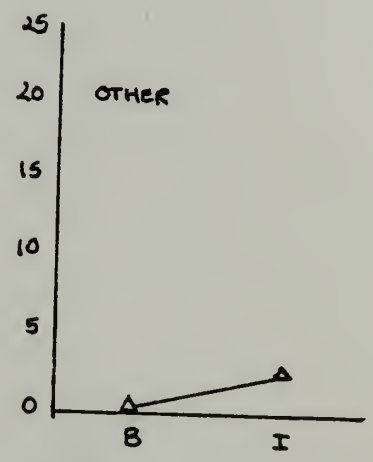


FIG 10: % OF INTERVALS OF SUBJECT B's 'OTHER' VOCALIZATIONS DURING BASELINE AND INTERVENTION.

setting the decrease in initiations and responses was an increase in ambiguous vocal behavior. This suggests two possibilities. Either Subject B emitted a greater number of vocalizations towards peers and adults during intervention which were unintelligible or he increased his rate of talking to himself as he played. The latter appears to be the most likely case when one considers his high percentage of solitary play. (See section on Play Behavior p.143).

"Other" vocalizations including laughs, smiles and play noises increased during intervention. No instances had been recording during baseline. These data are shown in Table 14 and Figure 10.

Table 14: Percentage of "other" vocalizations during baseline and intervention for Subject B.

Other	
Baseline	Intervention
0	2.3%

Although the increase appears quite sharp, it only accounts for 2.3% of all of Subject B's vocalizations.

Distribution. Subject B's initiations to adults and peers were well distributed across all persons in the classroom. Table 13 indicates the number of persons involved in

vocal interactions during baseline and intervention.

Table 15: Number of individual peers and adults involved in vocal interaction with Subject B.

	Baseline		Intervention	
	Initiations	Responses	Initiations	Responses
To Adult	7	6	7	4
To Peer	10(2sp, 3nsp)*	1 (1nsp)	6(2sp, 4nsp)	0

* sp = special needs student

nsp = non special needs student

Of the seven adults, two were more frequently involved than the remaining five. They were the classroom teacher and the teacher's aide and they received 10 and 9 initiations respectively from Subject B. The frequency of incidents of the other classes were evenly distributed.

Responses, unlike initiations, differed considerably in distribution between adults and peers. In fact, only one peer was involved in a vocal interaction requiring a response from Subject B. This discrepancy suggests that Subject B was rarely addressed by fellow students during the free play period.

Content Analysis. For Subject B, the content analysis was conducted on a sample of approximately 50% of his free play vocalizations. The limited sample represents the vocalizations that were accurately recorded. Subject B has several speech problems including poor articulation and word pronunciation. He also displays a limited repertoire of expressions but several were identified throughout the course of the investigation.

Of the data collected, it appeared that the major function of Subject B's initiations were to attract attention of a listener. For example, he used phrases like "Hey you" or would call a person by name, particularly the classroom teacher. Other frequently distinguished expressions included "What's that?", "Come here.", and "Look". These phrases would all fall into the category of mands as they function as requests and are controlled by the presence of a listener. Other verbal operants emitted by Subject B included tacts and intraverbals. Tacts generally occurred during object play when Subject B would name the toys or men he was playing with. For example, when playing with a "Batman" doll, he repeatedly said "Batman".

Table 16 lists the frequencies of various vocalizations and their class of verbal behavior. The data represent 21 days of observation--11 of baseline and 10 of intervention.

Table 16: Frequency and classification of Subject B's vocalizations.

Subject B's Vocal Verbal Behavior	Classification	Frequency B	Frequency I
Greetings: "Hi" or "Hey you"	Mand/Tact	8	8
Calling an Adult by Name	Mand/Tact	4	4
"Look"	Mand	2	3
"Come Here"	Mand	2	3
"More"	Mand	2	0
"What's that"	Mand	4	1
"Watch this"	Mand	0	1
"No"	Intraverbal	1	1
"Yah"	Intraverbal	1	1
Naming Objects (e.g., Batman, Spiderman)	Tact	1	2
	Total	25	24

Non-Special Needs Children. Comparisons between both special needs subjects and their nonhandicapped peers, revealed marked differences in types and frequencies of vocalizations. The combined data for six non-special needs students are presented in Tables 17 and 18. These data represent 96 total minutes of observation in contrast with 235 minutes for Subject A and 220 minutes for Subject B. They

are transformed to indicate the mean percentage of occurrence of the behaviors for the group. The group mean is an average of individual student means. (The raw frequency data for these classes of behavior can be found in Appendix IV.)

Table 17: Percentages of "initiations" and "responses" of non-special needs students

	Non-Special Needs Students	
	Initiations	Responses
To Adult	7.0%	6.5%
To Peer	15.8%	12.2%

Table 18: Percentages of "Ambiguous" and "Other" classes of vocal behavior of non-special needs students

Non-Special Needs Students	
Ambiguous	Other
12.7%	3.4%

The most striking differences in the special and non-special students' vocalizations can be seen in the percentage of time engaged in verbal interactions. Vocalizations by the nonhandicapped children occurred in 57.6% of all intervals whereas they only accounted for 6.5% of Subject A's

intervals during intervention and 35.3% of Subject B's. Even though the figure for Subject B approaches that of the nonhandicapped, it should be noted that 23% is accounted for by "Ambiguous" vocalizations.

Looking more closely, data on initiations and responses to peers indicate that there is a discrepancy in the verbal repertoires of these two groups. For example, the target subjects initiated vocalizations during 3.4% and 1.5% (Subjects A & B, respectively) of the intervals as compared to 15.8% for the non-special needs students. Similar differences were seen for responses to the vocalizations of others. Subjects A and B responded to others during .6% and 0% of the intervals in comparison with 12.2% for the nonhandicapped children. It seems clear that normal preschool children spend a significant portion of time in verbal interactions with their peers.

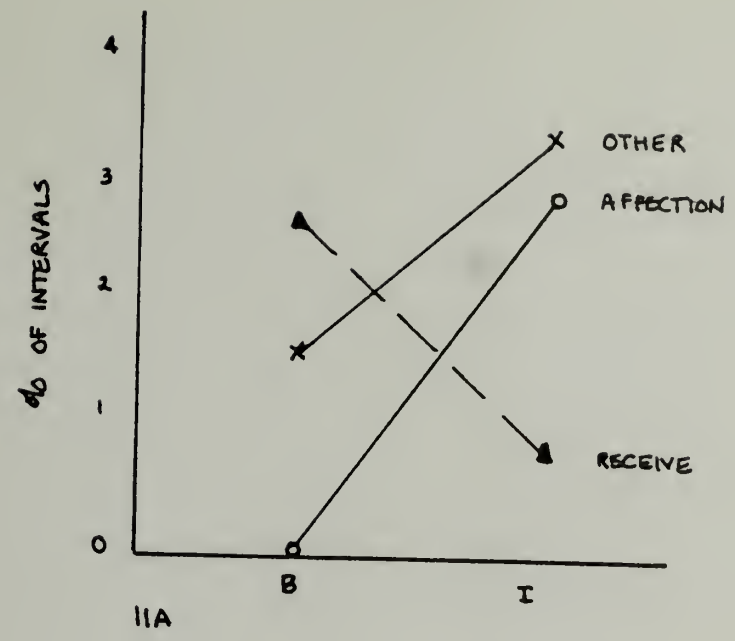
Another interesting outcome of the observations of the normal students was the high percentage of intervals of ambiguous verbal behavior. It will be recalled that ambiguous vocalizations were scores in the event that subjects talked to themselves or narrated their play activities, when they supplied sound effects during play, or if it could not be determined to whom a statement was directed. Judging from the written notes on this category, recorded during observations, the majority of the ambiguous vocalizations resulted

from play narration and sound effects.

Social Interactions.

Subject A. A summary of the data for Subject A's social interactions can be found in Table 19 and Figures 11a and 11b. Shown are the percentages of occurrence of each of the seven behaviors scored during free play. Raw frequency data and daily percentage can be found in Appendix .

As can be seen from the Figures, only two of the seven behaviors, "affectionate" and "other" increased following intervention. The other behaviors (Take-tug-grab, give, offer, and aggression) remained stable with the exception of "receive" which decreased. The decrease and two incidents of increased percentages coincide with the observed changes in Subject A's play behavior. (Data on play behavior are presented below.) During baseline, Subject A engaged in a high percentage of solitary and parallel toy play. Often, this would involve play at an activity table (e.g., sand or water tables) where a number of objects were available. Under these conditions, children would frequently exchange toys or objects which might lead to a higher proportion of "receive" intervals than a different mode of play (e.g., a gross motor activity such as swinging or climbing the ladder). Following intervention, Subject A



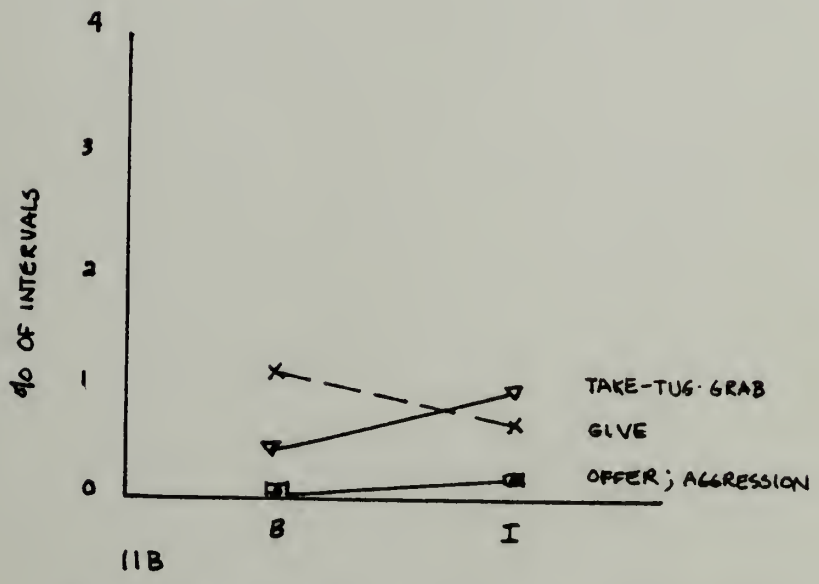


FIG. 11: PERCENT OF INTERVALS OF SOCIAL INTERACTION BEHAVIORS DURING BASELINE AND INTERVENTION FOR SUBJECT A.

Table 19: Percentages of Subject A's social interaction behaviors during baseline and intervention.

	Baseline	Intervention
Give	.94%	.58%
Offer	0	.1%
Affection	0	3.0%
Aggressive	0	.1%
Receive	2.8%	.8%
Take-Tug-Grab	.3%	.7%
Other	1.6%	3.5%

was observed to play cooperatively more often than either solitarily or in parallel and her time was spent primarily at the climbing structure. Here, in addition to gross motor activity, she emitted many affectionate behaviors such as patting and putting her arm around a peer while going down the slide. As this type of play generally did not involve toys, it lessened the chances that materials would be exchanged and "receives" would occur. The increase in "other" behavior during intervention represents a change in the frequency of social interactions with Subject A that were initiated by other peers.

Distribution. The percentages of social interactions involving adults and peers were compared during baseline and

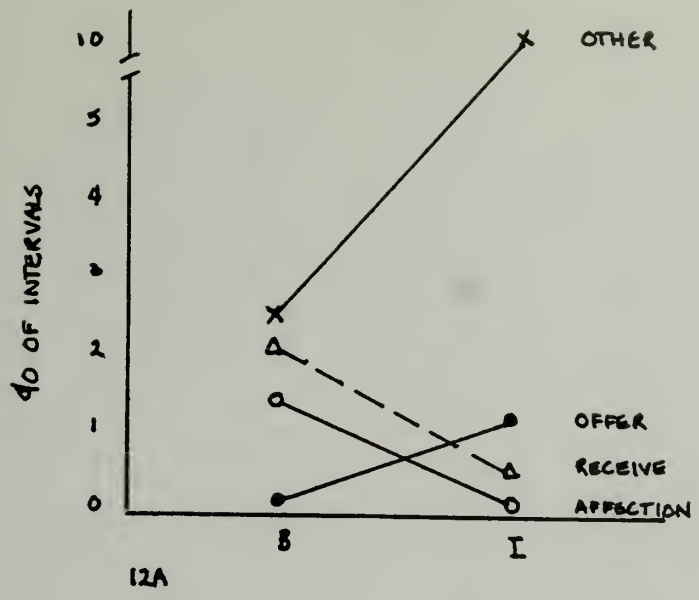
intervention. These data appear in Table 20.

Table 20: Subject A's percentages of social interactions with peers and adults during baseline and intervention

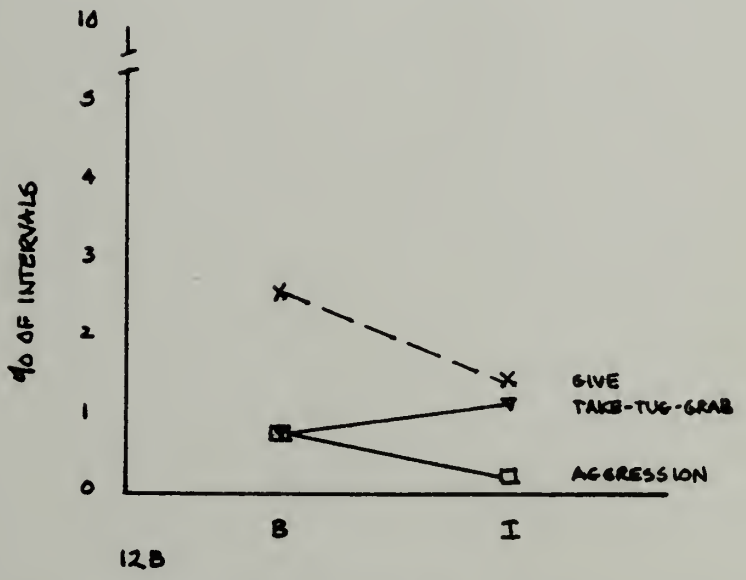
	Baseline	Intervention
To Adult	4.68%	2.78%
To Peer	3.44%	8.42%
Combined Adult & Peer	8.12%	11.2%

The overall percentage of social interactions (combined adult & peer) increased from 8.12% to 11.2%. These figures represent a substantial increase in peer interactions and a slight decrease in interactions with adults. Within the peer group, it was observed that the distribution of individuals with whom interactions occurred also grew. During baseline, Subject A was observed to interact with five non-special needs students whereas after intervention she interacted socially with ten peers. The ten students were four special needs children and six non-special needs. The distribution of adults changed marginally from three during baseline to two during intervention.

Subject B. The percentages of social interactions for Subject B appear in Table 21 and Figures 12 a and 12b.



12A



12B

FIG 12: PERCENT OF INTERVALS OF SOCIAL INTERACTION BEHAVIORS DURING BASELINE AND INTERVENTION FOR SUBJECT B.

Table 21: Percentages of Subject B's social interactions during baseline and intervention.

	Baseline	Intervention
Give	2.6%	1.3%
Offer	0	1.1%
Affection	1.47%	0
Aggression	.7%	.2%
Receive	2.2%	.2%
Take-Tug-Grab	.7%	1.1%
Other	2.6%	10.5%

Percentages for all classes of behavior except "other" remained relatively stable over the baseline and intervention periods. There were slight decreases of approximately 1% for the behaviors "give" and "receive". "Other" social interactions, in contrast, showed a substantial increase over the baseline level. As was the case with Subject A, the behaviors represented included social interactions initiated by peers and assistance from both peers and adults during play activities.

Distribution. The distribution of social interaction percentages during free play involving peers and adults appear in Table 22.

Table 22: Subject B's percentages of social interactions with peers and adults during baseline and intervention.

	Baseline	Intervention
To Adult	5.4%	4.2%
To Peer	3.6%	6.6%
Combined Adult & Peer	9.06%	10.8%

The overall percentage of social behavior increased from 9.06% to 10.8%. This figure represents an increase in peer directed social interactions and a slight decrease in adult centered interactions.

The distribution of interactions among individual peers demonstrated several changes. During baseline, Subject B interacted with 7 fellow students, 3 special needs and 4 non-special needs children. This figure rose to 9 students during intervention representing 3 special needs and 6 non-special needs students. The distribution of adults, conversely, diminished slightly from 6 individuals during baseline to 4 during intervention.

Non-special Needs Children. Of the 24 observations over 9 days of observation of non-special needs subjects totalling 96 minutes, only 14 social interactions were observed for an overall percentage of 3.61%. The percentages

of intervals of the individual categories are presented in Table 23.

Table 23: Percentages of social interactions of non-special needs subjects.

	Non-Special Needs
Give	1.5%
Offer	0
Affection	.9%
Aggression	.3%
Receive	.63%
Take-Tug-Grab	.9%
Other	0

The percentages of social interactions emitted by this group of children was surprisingly low. In fact, they were somewhat lower than the rates and percentages obtained for the two special needs subjects.

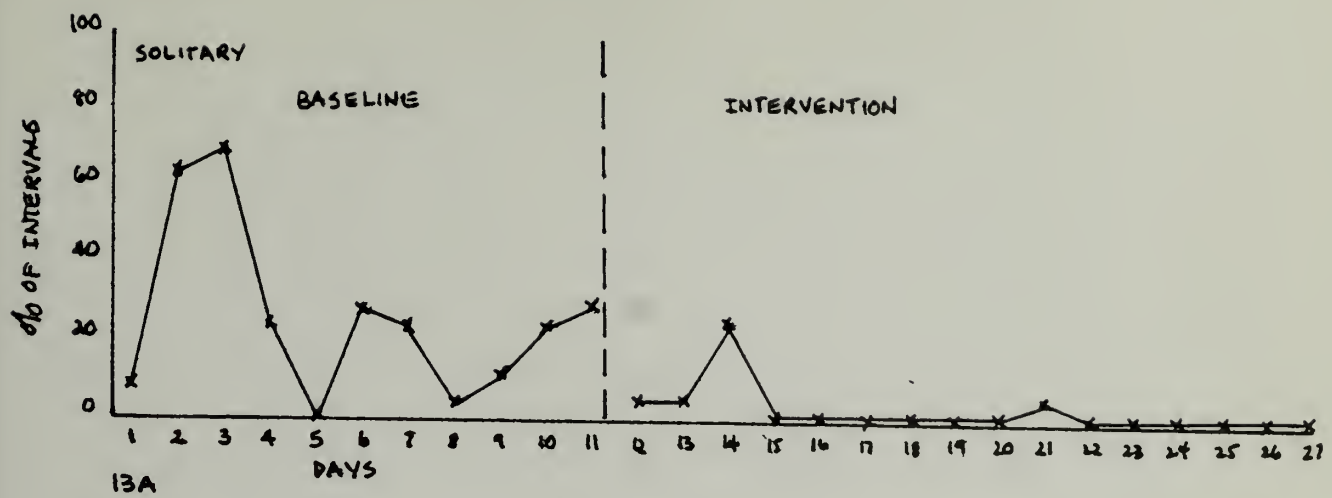
Play Behavior.

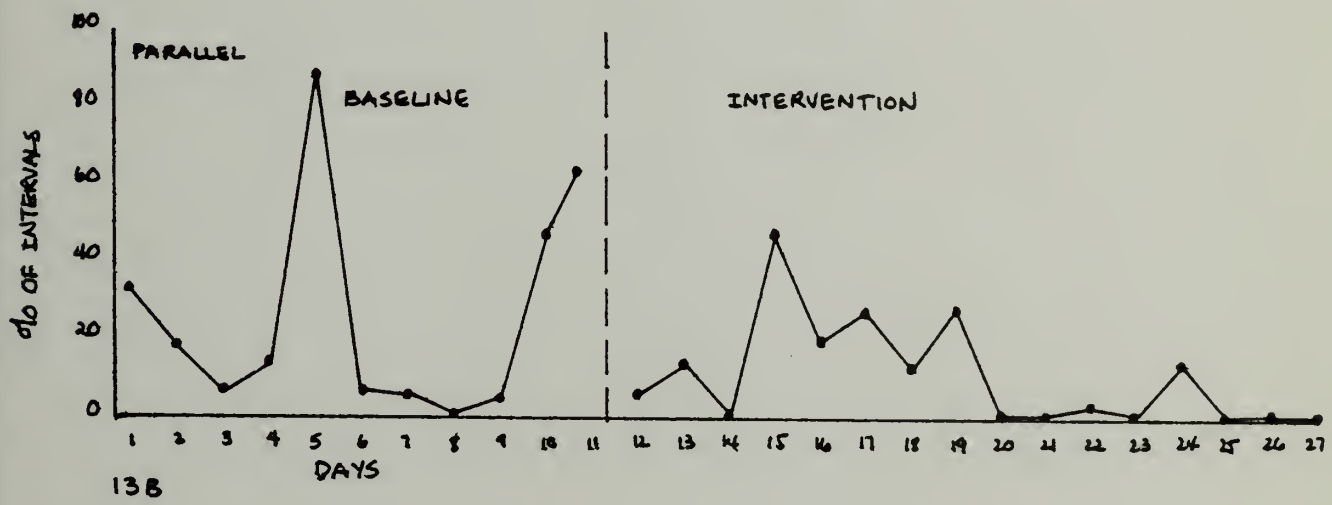
Subject A. The daily percentages of solitary, parallel, and cooperative play for baseline and intervention periods are shown in Figures 13a, b, and c. The baseline period was dominated by solitary play where the mean percentage of intervals was 28% and parallel play accounting for 22% of the

observed intervals. In contrast, Subject A engaged in cooperative play an average of 8% of the baseline period. The mean percentages of play behaviors appear in Figure 14. Following the introduction of the intervention, the baseline trend in the data was reversed. Solitary play dropped to near zero levels, while cooperative play increased to a mean of 42%. Parallel play showed a decrease, particularly towards the end of the intervention phase, from a mean of 22% to 10%. Although not evident in these figures, anecdotal records of activities during play indicated that the quality of play shifted along with the frequency.

Figure 15a presents the data for the fourth type of play behavior recorded, "passive watch". Like solitary play, percentages of passive watch are considerably higher during baseline than intervention. This is not surprising considering the complimentary nature of the two behaviors. When engaged in solitary play, Subject A would readily attend visually to the activities around her. During cooperative play, however, she engaged in more gross motor activities and was less likely to respond to other events.

The final two behaviors in this category, "with adult" and "other" are shown in Figures 15b and 15c. There were no apparent changes in either variables as a function of the intervention program. The importance of these





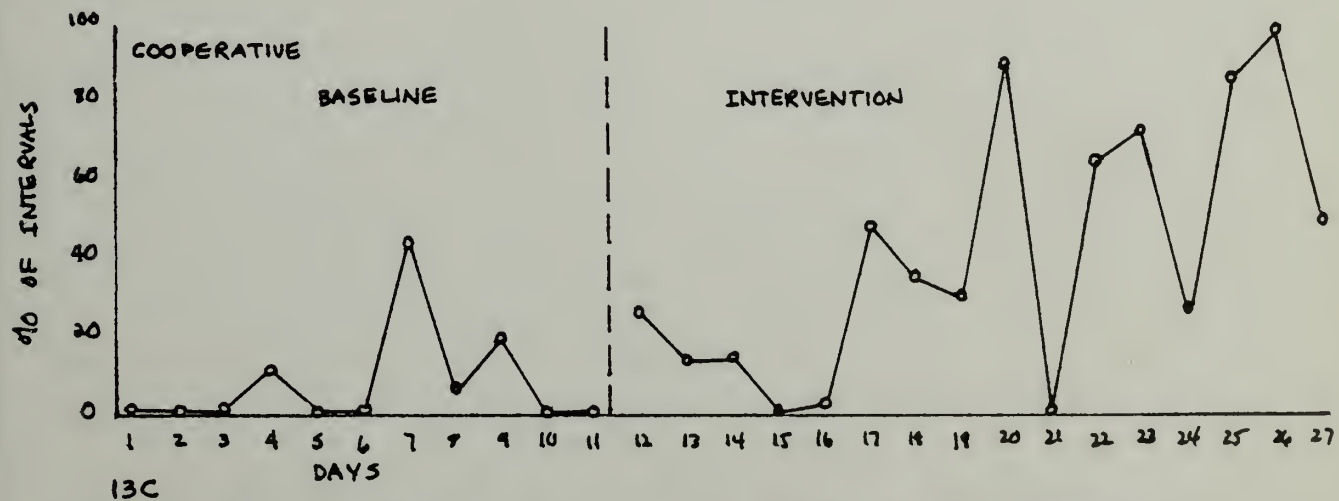


FIG 13: PERCENT OF INTERVALS OF SOLITARY (A), PARALLEL (B), AND COOPERATNG (C) PLAY BEHAVIOR DURING BASELINE AND INTERVENTION FOR SUBJECT A.

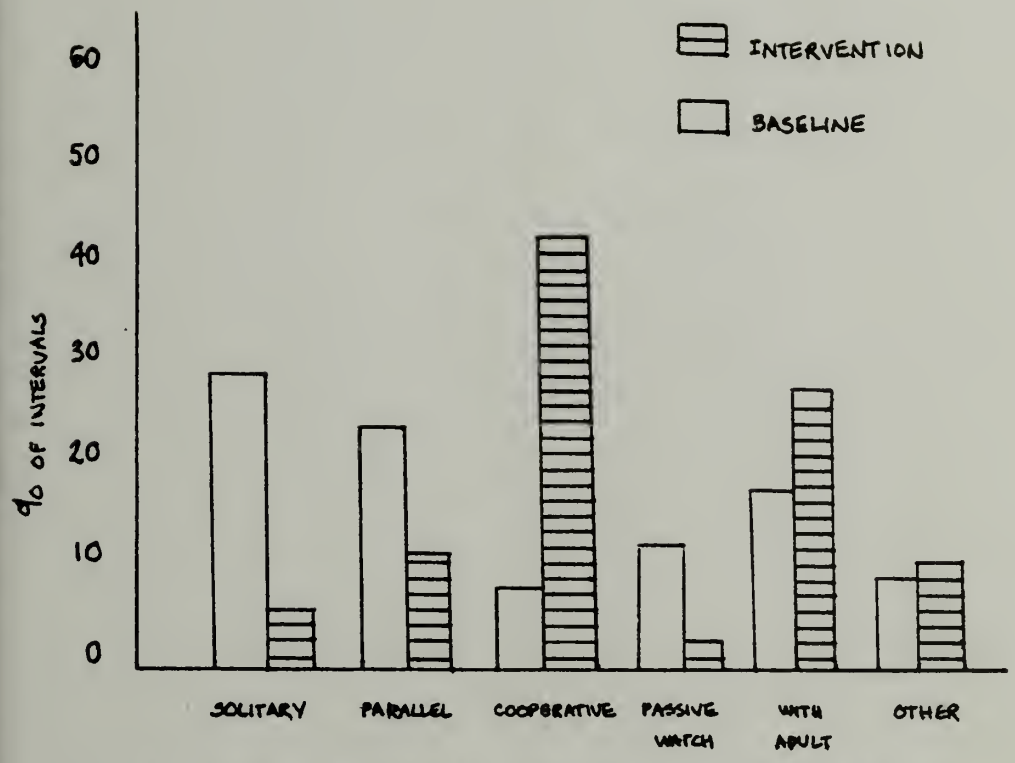
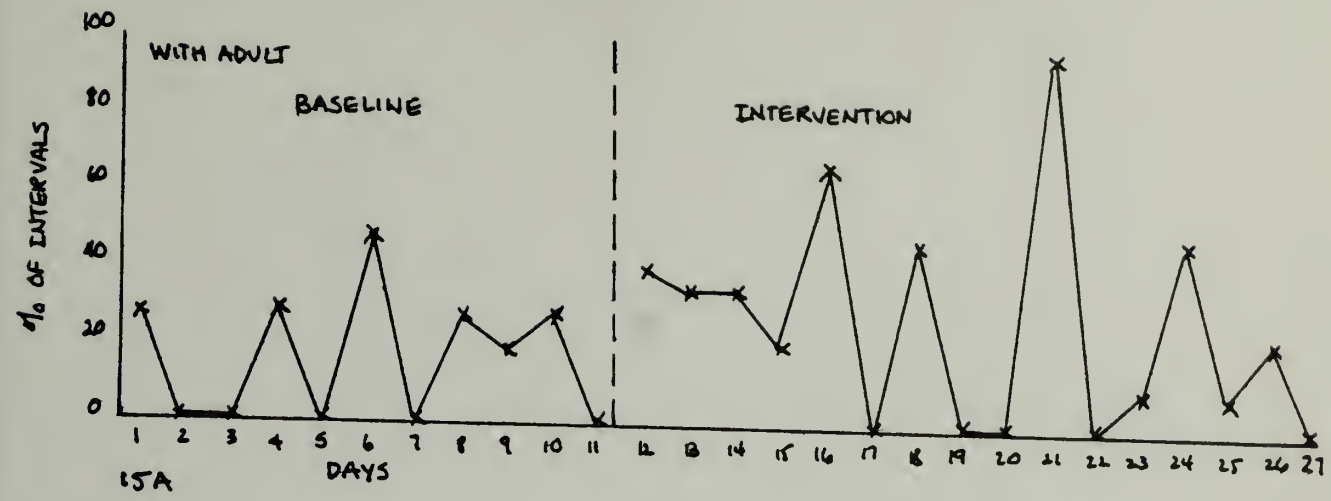


FIG 14: MEAN PERCENTAGES OF PLAY BEHAVIOR DURING BASELINE AND INTERVENTION PHASES FOR SUBJECT A.



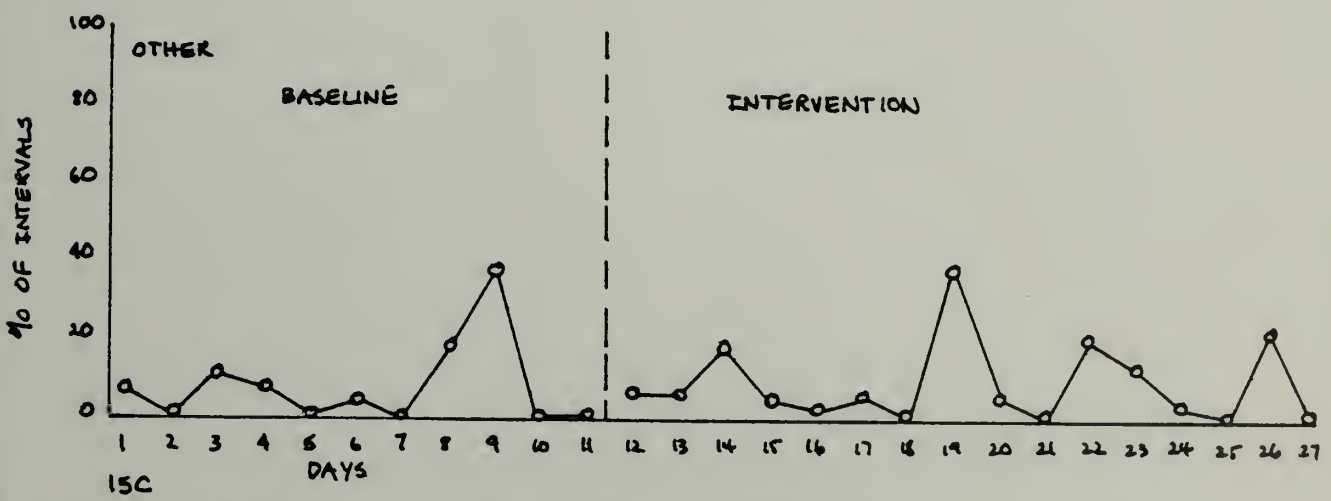


FIG 15: PERCENT OF INTERVALS OF WITH ADULT (A), PASSIVE WATCH (B), AND OTHER (C), PLAY BEHAVIOR DURING BASELINE AND INTERVENTION FOR SUBJECT A.

measures, though, is in the documentation of Subject A's behavior when she was not involved in one of the four behaviors presented above.

The "with adult" category is even more difficult to interpret as it is not an independent behavior. The incidents recorded include both those initiated by the subject and those initiated by the adults. Therefore, it cannot be used as a measure of student seeking adult interactions. "Other" behavior indicates the degree with which the subject changed activities or wandered around the room. The no-change status of this measure suggests that Subject A remained at her respective activities during baseline and intervention for approximately the same length of time.

Distribution. The specific individuals with whom Subject A interacted cooperatively were compared for baseline and intervention periods. In the former, she was observed to play with 4 different students (1 special needs and 3 non-special needs). This figure rose to 10 students during intervention representing 3 special needs and 7 non-special needs peers.

Subject B. The daily percentages of solitary, parallel, and cooperative play are shown in Figures 16a, b, and c. A bar graph, Figure 17 presents the mean percentages of the play behaviors during baseline and intervention. The base-

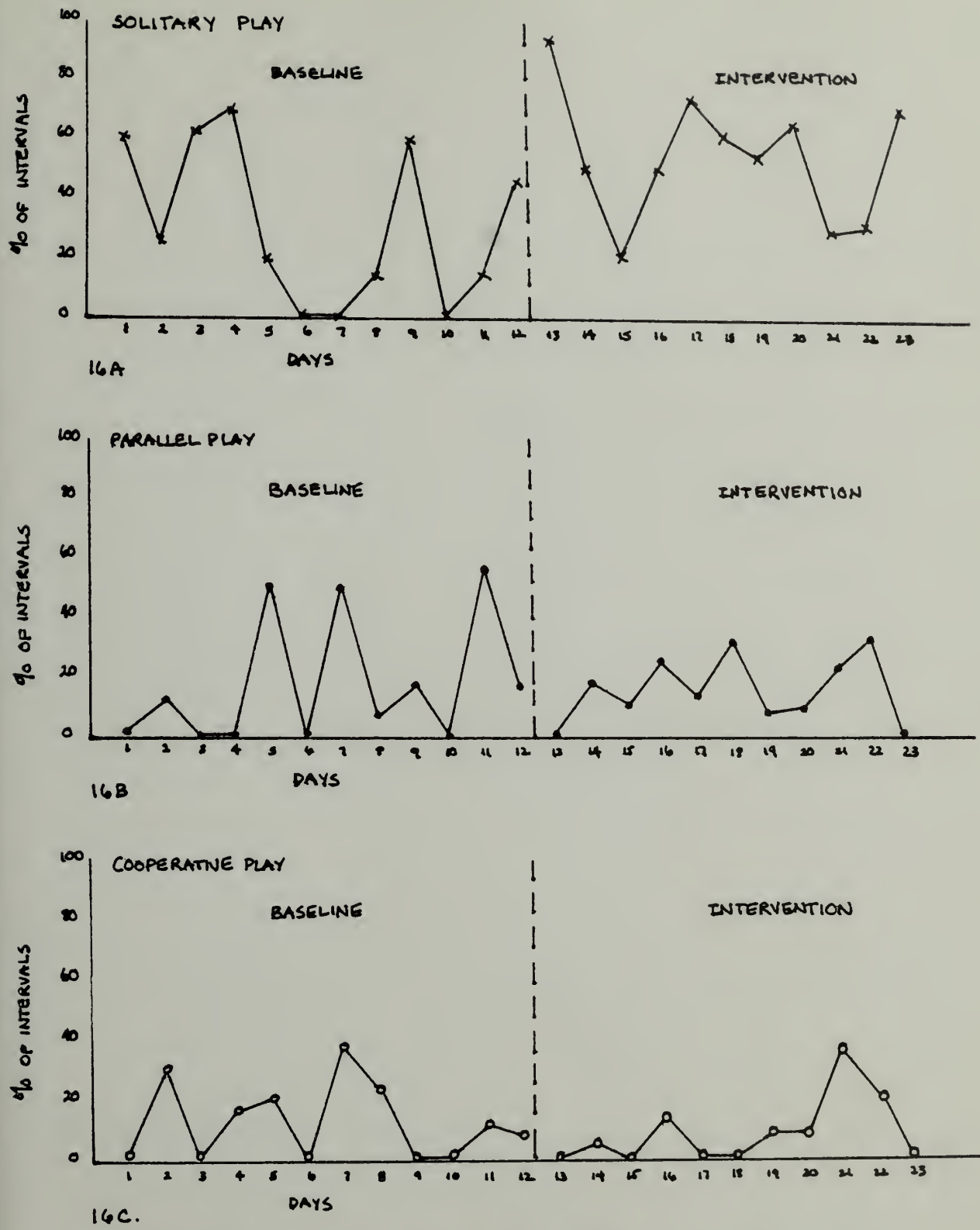


FIG 16: PERCENT OF INTERVALS OF SOLITARY (A), PARALLEL (B) AND COOPERATIVE (C) PLAY DURING BASELINE AND INTERVENTION FOR SUBJECT B.

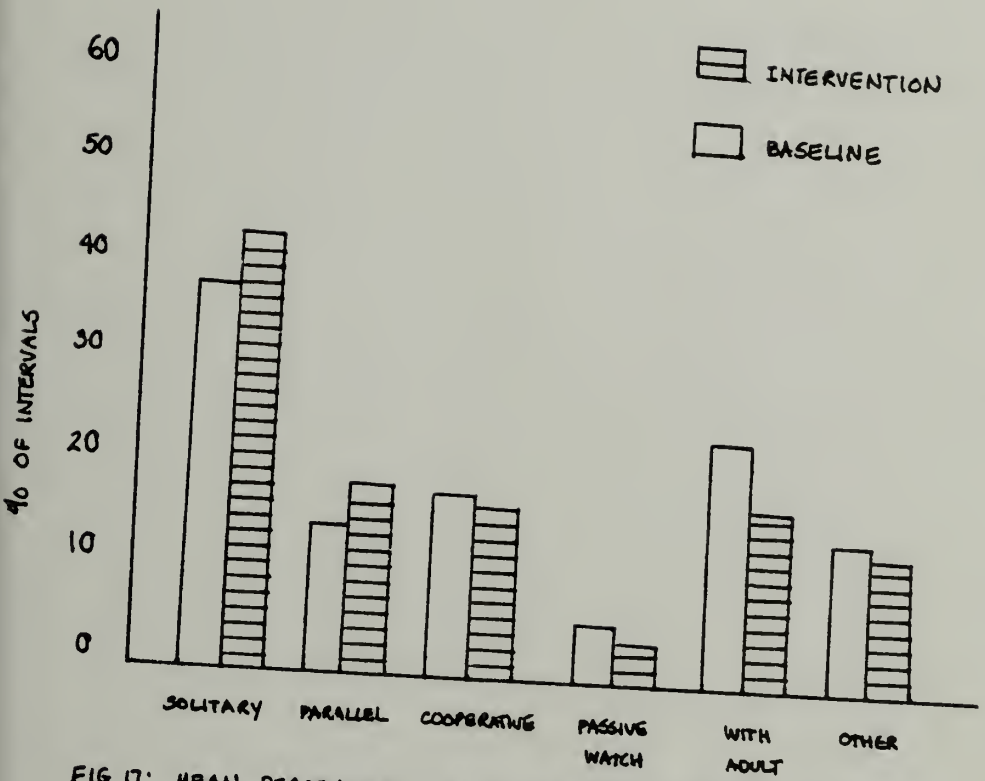


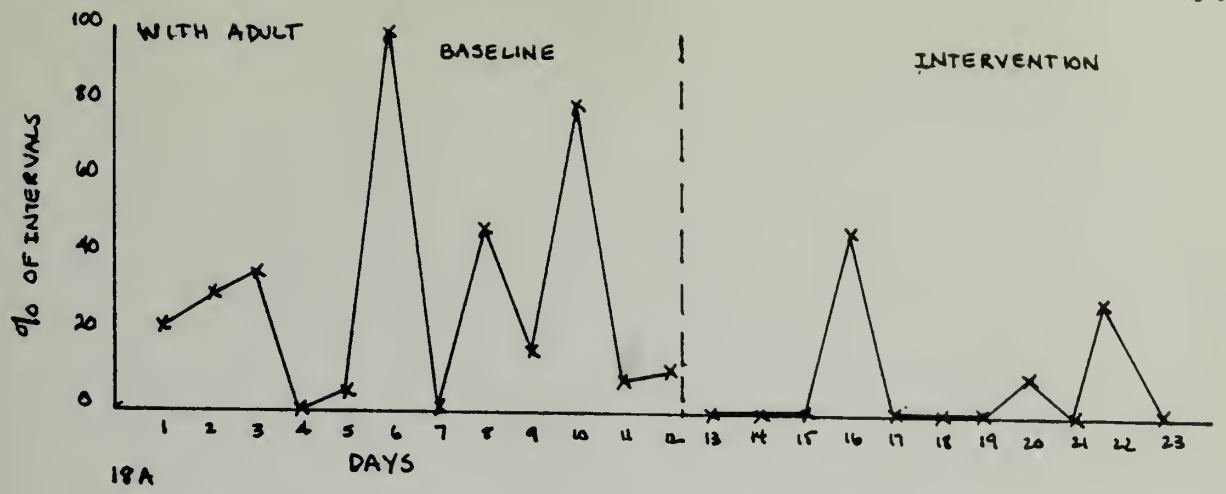
FIG 17: MEAN PERCENTAGES OF PLAY BEHAVIOR DURING BASELINE AND INTERVENTION PHASES FOR SUBJECT B.

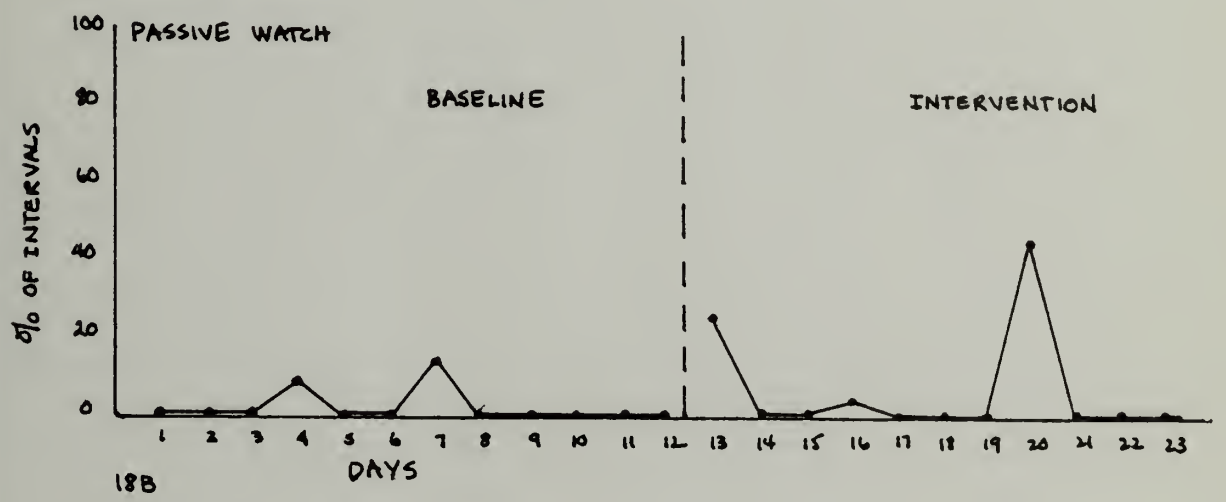
line levels of Subject B's play showed solitary play as the prevalent form followed by parallel and finally cooperative play. This trend was similar to the baseline performance of Subject A. However, the intervention phases contrasted sharply. After the introduction of the social interaction game, the percentages of Subject B's solitary and parallel play increased. A minor drop in the level of cooperative play also occurred.

Figures 18a, b, and c show the daily percentages of the remaining three behaviors in this category, Passive Watch, With Adult, and Other. There were no apparent systematic changes in these behaviors as a function of the intervention, although, they are useful data as they document Subject B's activity during free play.

Distribution. Despite the fact that the frequency of cooperative play did not increase during intervention, there was a slight change in the number of peers involved in cooperative interactions. During baseline, Subject B was observed in cooperative play with 5 students (1 special needs and 4 non-special needs). This figure increased to 8 students (3 special needs and 5 non-special needs) during the intervention phase.

Non-special Needs Subjects. The six variables in the play behavior category are summarized in Figures 19 and 20 for the nonhandicapped children. Figure 19 shows the per-





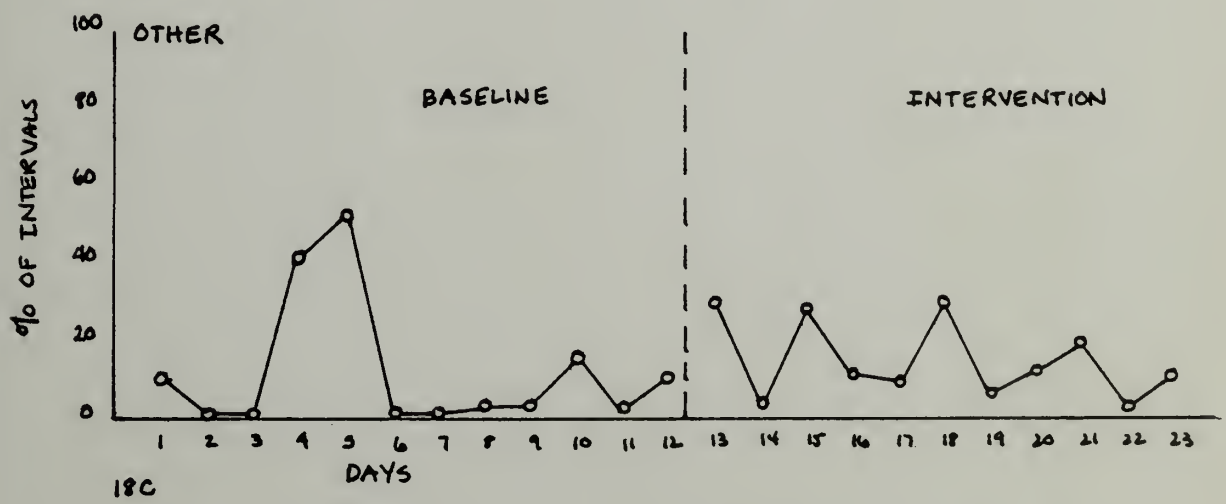


FIG 18: PERCENT OF INTERVALS OF WITH ADULT (A), PASSIVE WATCH (B), AND OTHER (C) PLAY BEHAVIOR DURING BASELINE AND INTERVENTION FOR SUBJECT B.

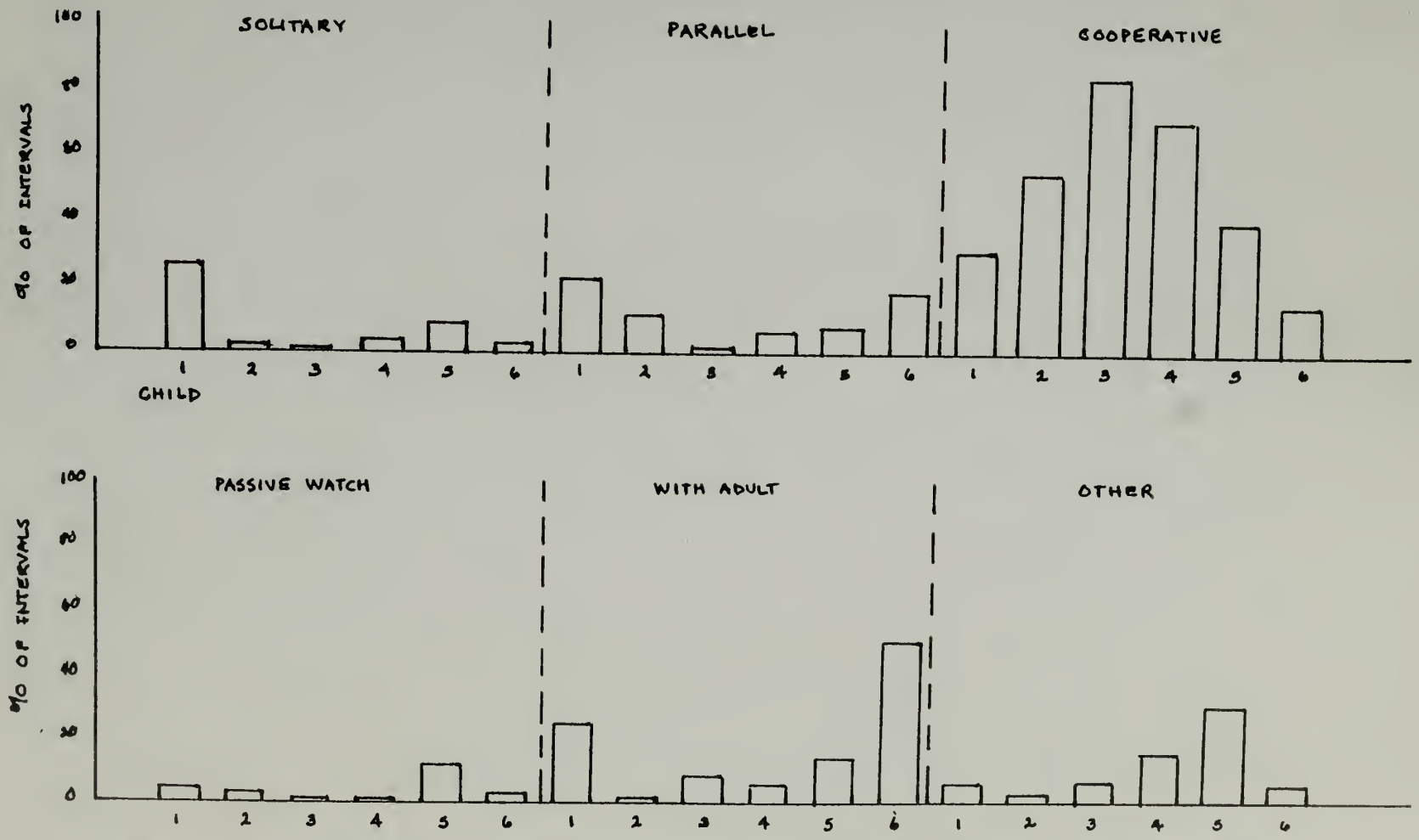


FIG 19: MEAN PERCENTAGE OF PLAY BEHAVIOR OF INDIVIDUAL NON SPECIAL NEEDS STUDENTS. (NUMBERS UNDER THE BARS REPRESENT INDIVIDUAL STUDENTS.)

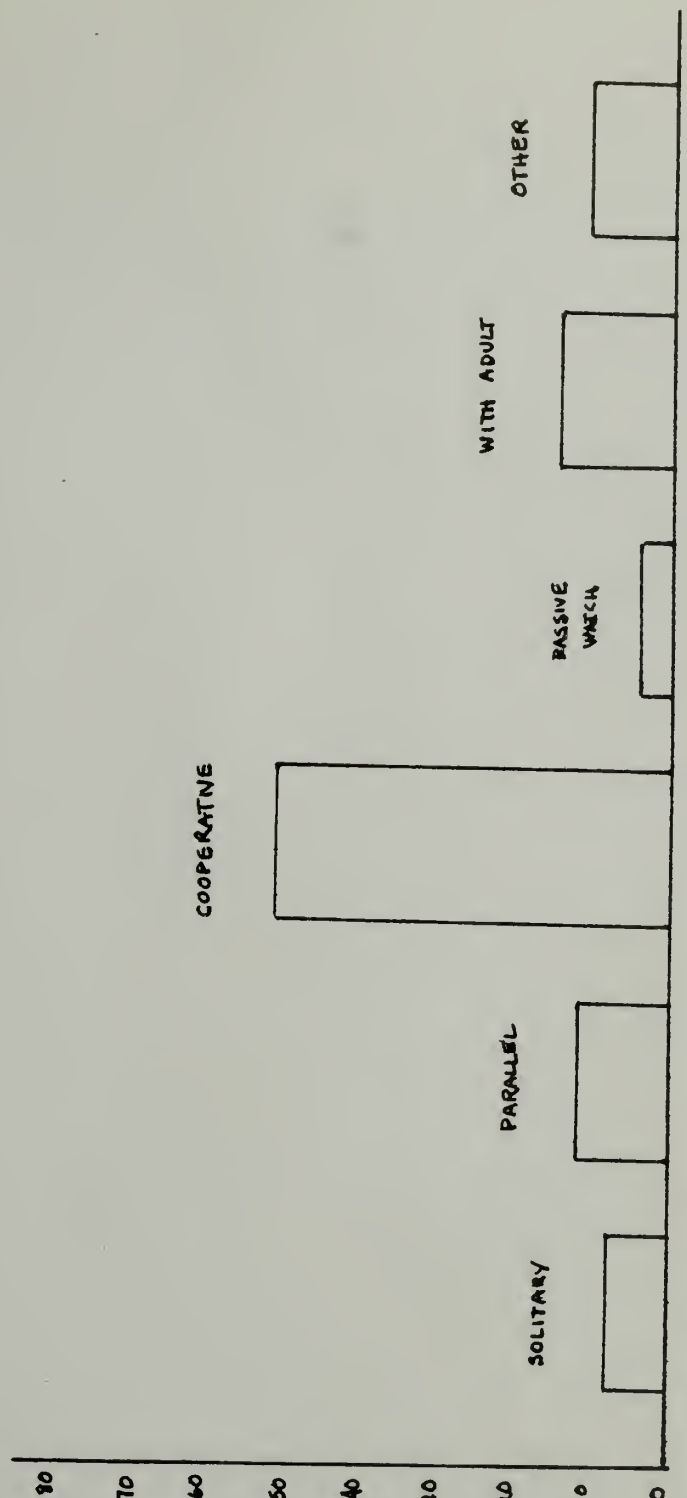


FIG 20: MEAN PERCENTAGE OF PLAY BEHAVIOR OF NONSPECIAL NEEDS STUDENTS.

centages for individual children, whereas these data are combined for a group total in Figure 20. As can be seen, cooperative play was by far the dominant form of play behavior. It accounted for nearly 55% of the free play activities of this group. Solitary and parallel play, on the other hand, involved less than 10% and 12% respectively.

In comparing the play profiles of the two target subjects to the nonhandicapped subjects, Subject A's performance during intervention most closely approximates the pattern of the nonhandicapped group. Subject B's profile differs primarily in the degree to which solitary play occurs.

CHAPTER V

DISCUSSION

The present study explored the impact of a social interaction game on the direct and collateral social behavior of handicapped preschoolers in a mainstreamed setting. The results obtained from individual target subjects and non-handicapped students are open to several interpretations depending on the scope of the analysis. Three levels of analysis were considered for the present data, each one representing an increase in breadth.

The first level focuses on the clinical effectiveness of the intervention. The major issues at this level concerned changes in the behavioral repertoires of the special needs students as a result of the intervention program. The importance of this level is based on both ethical and practical considerations. For example, permission to participate in the study was granted by parents, the school system, and the Early Education Center staff, with the expectation that the project would potentially benefit the students. To meet the commitment to parents, teachers, and the clients themselves, the investigator has the responsibility to determine the program's effect on an individual basis.

The adequacy of the methodology was the central concern of the second level of analysis. At this point, the generalizability of procedures are considered. Questions are considered such as whether or not the same intervention could be applied efficiently and effectively to another group of students in a different setting. What needs to be determined are the factors present in the current setting that contributed to the outcome. The final perspective takes a broader look at the problems of evaluating mainstreaming programs. Although similar to the concerns of the second level of analysis, the focus here is predominantly on how the effects of integration can be assessed given variety of sizes, composition, and facilities of mainstreamed classrooms. The difficulty in conducting such evaluations lies partly in the recency of mainstreaming as an intervention technique. Educators and evaluators are unsure of what positive and/or negative effects can be expected from mainstreaming settings. Part of the problem is the number of factors influencing program effectiveness. Children with a wide array of handicaps, needing a variety of special services are placed in different sized classes, with different ratios of handicapped to nonhandicapped students. Any one or combinations of these factors may be the most salient variables in determining the impact of the program.

A related issue concerns the definition of effective integration. Mainstreaming, used in its most general sense, may simply define a placement strategy. That is, handicapped and nonhandicapped students are grouped together in one classroom. However, mainstreaming can also represent an active attempt by teachers to adapt teaching methods to the special needs students while accomodating the needs of all of the students. Unfortunately there is no ideal system of accomplishing mainstreaming. Thus, because instruction varies so much from classroom to classroom, there is no single appropriate method for assessing its effectiveness. At the present time, therefore, it makes more sense to focus on individual classes in which normal and handicapped children are integrated. In the future, it may be possible to combine the data from a large number of individual studies to begin to assess some of the more general effects of integration. Some variables that may be considered include the ratio of handicapped to nonhandicapped students, the severity of the handicaps, and teacher training and instructional support services to name a few. In the meantime, evaluation methods are needed that can be adapted to the conditions in any given classroom. As researchers, we must be able to evaluate mainstreaming programs despite their diversity and the

constraints imposed by uncontrollable sources of variability.

Within the final level of analysis the question of classroom ecology must also be addressed. In particular, this issue concerns the effects of integration on the non-handicapped students as well as on the handicapped. Traditional methods of determining program impact is to study its effects on the client population for whom it was designed. However, this method fails to take into account any concommitant changes or effects on the normal children also enrolled in the program. Even if the program benefits the special needs student, it cannot be considered successful if services for the regular education students are compromised. It is to these issues that the final level of analysis is directed.

Clinical Effectiveness.

The impact of a clinical treatment procedure needs to be considered in terms of the impact on the individual client. Such an analysis is made possible using the single or within-subject experimental design. Traditional group comparisons common in psychological research overshadow treatment effects on individual subjects. Providing the investigator with an estimate of what the average behavior change would be for a certain population does not indicate which subjects

changed their behavior and in what direction the change occurred. Limitations of the group design for clinical practice have been discussed at length by advocates of single-subject methodology, including Sidman (1960), Chassan (1967), Bergin and Strupp (1972) and Hersen and Barlow (1976). The major criticism cited was that group mean scores cannot help the clinician determine which procedure will be effective for a particular client. The performance of subjects A and B, when reviewed simultaneously, illustrate this point. The behavior changes for the two special needs students reflected opposite trends. While there were increases in several classes of Subject A's behavior the same behavior decreased in Subject B. Had these two sets of data been averaged, one would have concluded that the intervention had no effect.

Subject A. In the case of Subject A, the introduction of the social interaction game resulted in increased frequencies of vocalizations and social interactions towards peers. Although the rates of these behaviors were still at marginal levels, the direction of change was positive. That is, her performance during intervention closely resembled the performance of the nonhandicapped students. The most dramatic change evidenced in Subject A's behavior was the shift from solitary and parallel play during baseline to cooperative play in the intervention period. The

mean percentages of cooperative play during the intervention phase equalled those demonstrated by the nonhandicapped students. These improvements in rates of behavior are interpreted as evidence of program success with this particular client.

In view of Subject A's entering behaviors, it is not surprising that positive results were obtained. It should be recalled that in relation to normal peers she began the program with depressed rates of social interactions--and was described as socially withdrawn. The most striking feature of her repertoire was the low rate of vocal verbal behavior, both initiations and responses. The behaviors were not the result of cognitive deficiencies as Subject A demonstrated an average level of intelligence on standardized tests of cognitive ability. The lack of vocalization was viewed as an inappropriately learned pattern of responding which was maintained by her environment. Apparently, "not talking" was reinforced by others who successfully interpreted her gestural communication. Both peers and adults were observed speaking for the subject when trying to understand her. Questions such as "You want (name of object)?" or "Do you want a turn?" were posed to elicit responses from her. This behavior on the part of peers and adults would certainly lessen the likelihood that she would initiate speech.

The features of the intervention that appear to have contributed to program effectiveness include both setting and process variables. For example, the games were played in a non-threatening setting where a single pair of students worked together in the presence of an adult. The game process provided the opportunity for subjects to earn tangible and social rewards from one another and the experimenter for correct vocalizations and social interactions. Players alternated turns giving each child an equal opportunity to approve verbal descriptions, give feedback, and offer the reinforcers. The repeated participation in the intervention program may have strengthened Subject A's verbal and social repertoires, thus enabling her to interact more appropriately during free-play.

The relatively nonintensive level of the intervention in terms of time commitment, frequency, and level of specificity of the requirements may also have contributed to the outcome. Here, they met with a peer and the experimenter twice a week for 15 to 20 minutes. For Subject A, this minimal amount of intervention produced positive changes in social, vocal, and play behaviors. This might not be true for other subjects with different entering behaviors (see results for Subject B). In other situations, perhaps, a more intense intervention program would be required. For example, it might be necessary to include

direct skills training before a child were able to participate.

Other factors that may have had a direct influence on the program's effectiveness include concomitant efforts by the classroom teacher and speech therapist to develop Subject A's verbal skills. Although such efforts were underway long before the introduction of the skill development game, the combination or procedures may have been critical in promoting change. Perhaps the combined efforts were sufficient to alter the conditions controlling Subject A's emission of verbal behavior. The notion of stimulus control suggests that altering her behavior requires a change in environmental stimulus cues and not just the acquisition of new skills.

The most significant changes in Subject A's social repertoire were the increases in positive affective behaviors: laughing, smiling, and affection. These changes are consistent with the literature on social interaction development among behaviorally handicapped children. In a theoretical analysis of social interaction studies, Strain and Shores (1977), presented findings on observational and experimental reports of handicapped children who were deficient in social skills. It was reported that handicapped preschoolers often have atypical social repertoires. That is, they differ in the frequency, duration or inten-

sity of responding. These deficiencies can be observed in children with a variety of disturbances including "severe developmental delay, emotional disturbance, sensory impairment, neurological impairment, and one or more physical handicaps" (Strain, Cooke, & Apolloni, 1976).

Three theories are currently available to explain the development of effective or social responses, the maturational, the cognitive-developmental, and the learning theory approach. Each has its own focus yet all recognize the importance of early experience on later social behavior. Of the three, only the learning theorists have attempted to identify relations between controlling stimuli and various social responses (Strain & Shores, 1977). Therefore, the learning theory research has been the most amenable to the formation of treatment strategies for clients lacking social skills. Traditional treatments within this approach, however, concentrated their efforts solely on the behaviors of the deficient child. They did not take into account the interactional nature of the social situation. Recent research on reciprocity has redirected the focus to the interactions of participants in a social encounter. Reciprocity, according to Patterson and Reid (1970), "describes dyadic interaction in which persons A and B reinforce each other at an equitable rate" (Patter-

son & Reid, 1979, p. 133).

The concept of reciprocity has a dual usage. Strain and Shores (1977) use it to explain the "development and maintenance of social responses which reflect an affective relationship" and as a "theoretical base for systematically developing empirically based educational programs to enhance the development of appropriate social behavior of young handicapped children" (Strain & Shores, 1977, p. 496). Looking at the changes in Subject A's behavior which occurred as a result of the intervention, one can appeal to the notion of reciprocity to account for the observed behaviors. The skill development game provided opportunities for reciprocal interactions. Each child had equal chances to provide positive responses. The importance of emitting positive responses was demonstrated in a series of observational studies by Hartup and his coworkers. They found that the amount of positive behavior emitted towards peers in a nursery school setting was positively correlated with positive responses received (Charlesworth & Hartup, 1967). The findings of the present study support their findings. As Subject A's skill in dispensing positive behaviors increased, so did her positive interactions with peers. Of particular interest was the increase in other social interactions as these data indicated increases in interactions initiated by peers.

Strain, Shores and Timm (1977) have demonstrated that withdrawn children's rates of social responses can be accelerated by providing appropriate positive social stimuli. Their study, conducted in a laboratory setting, used peer confederates to initiate social interactions with behaviorally handicapped students. Whether the findings would generalize to other settings was not established. It appeared, however, that the social stimuli did set the occasion for reciprocal responses. The findings of the study reported here support the notion of reciprocity and suggest that generalization across settings is possible. Although, the rates of behavior occurring in the new setting were not necessarily as high as those reported in studies in the lab setting.

A final point noted by Strain and Shores (1977) regarding reciprocal interactions among preschoolers is that appropriate language skills appear to be a necessary condition for reciprocity to occur. Without similar verbal skills, it is unlikely that a pair of children can achieve equality in positive reciprocal interactions. Language may then play a very important role in determining the level of reciprocity achieved in a relationship between handicapped and nonhandicapped children. Although, it is recognized that language skills are not the only behaviors involved in reciprocal interactions, they appear to be of

central importance (Strain, Shores, & Kerr, 1976). As previously noted, Subject A's language difficulties were probably a function of poor stimulus control rather than a lack of comprehension or knowledge of specific vocabulary.

As can be seen in the functional analysis of Subject A's vocalizations her repertoire included the basic classes of verbal operants--mand, tact, echoic, and intraverbal. Although her rate of vocal verbal behavior increased slightly, its low rate might cause one to question whether Subject A and nonhandicapped peers shared similar language repertoires. This brings up two points. First, it should be noted that Skinner's analysis of verbal behavior includes non-vocal responses. In Subject A's case, she possessed the basic operants which could or might not have been emitted vocally. Therefore, it is likely that a good deal of language based interactions occurred with Subject A responding in a non-vocal fashion. A second point to consider is that equivalence in language repertoire is a loosely applied concept. Perhaps it is sufficient if both parties have overlapping repertoires and are not matched exactly on both comprehensive and oral speech functions. In Subject A's case, the presence of the different verbal operants and her general success in following classroom procedures etc. suggested that she

has considerable language skills. Now in turning to the results of Subject B, one finds a different set of outcome data showing opposite trends from Subject A.

Subject B. The origins of Subject B's handicaps were notably different than Subject A's. He was a Down's Syndrome child who exhibited poor language skills, cognitive deficits, and developmental delays. He scored well below average for his age on standardized tests of cognitive ability. Despite these differences in etiology between Subject B and A, both experienced similar functional handicaps. For example, both children rarely used vocal verbal behavior to communicate, and thus engaged in little interaction with their peers. Considering the similarities one might have expected success with the same intervention program. The results of the present study, however, indicate that such an assumption cannot be supported. Nevertheless, the findings are extremely informative and provide a basis for developing future treatment programs for handicapped children.

Subject B's performance during intervention showed relatively little change in rates of vocalizations and social interactions to peers and adults. In fact, the only improvements were in "other" behavior which suggests that the peer interactions structured by the game positively effected interactions initiated by peers. For example,

if Subject B was having difficulty completing a task another student might offer assistance which would be recorded as "other" in the social interaction category. Although, it should be noted that assistance was the main form of peer interactions. The other children would frequently come to the aid of Subject B, when he encountered difficulty with manipulative tasks.

In retrospect, the limited generalization of results for Subject B do not seem surprising. Reviewing his entering repertoires, it seems likely that his developmental deficiencies were more severe than could be ameliorated effectively by the skill development game. As discussed in the previous section, the intervention was designed to minimize intrusiveness. Perhaps, given Subject B's special needs, it was not structured sufficiently enough to change behaviors in the free-play setting. As it was, only slight behavior changes were observed during game sessions themselves. For example, experimenter prompts were required in all sessions although they diminished during the final two intervention periods. Prompts had been included to maintain the sequence of steps in the game and to elicit the verbal description of the picture cards and were utilized to cur 78% of the verbalizations with Subject B. (In contrast, it was never necessary to prompt Subject A's verbalizations.)

Sample prompts for Subject B included questions like "What is on the card?" or "What is the boy doing?" or even naming an object on the card. Spontaneous descriptions did occur in the final two games, accounting for 2 of the 5 trials in each. It was necessary to prompt the game sequence for both subjects yet again the percentages varied considerably. It was necessary to prompt Subject A to follow the game sequence on only 42% of the trials in comparison with 83% for Subject B. Despite the fact that the picture cards began to exert more stimulus control by the end of the intervention phase, it is clear that the situational cues controlled Subject A's behavior to a much greater degree than Subject B.

Why the game stimuli were ineffective can be explained by considering both cognitive deficits and attentional problems. Subject B's behavior appeared to be under the control of irrelevant stimuli in the experimental setting. In other words, he would be described as highly distractible with a low attention span. The problem seen in following the sequence would appear to be a learning problem. He failed to acquire the chain of responses necessary for the game, yet could follow them if prompted. Subject A on the other hand, appeared to have mastered the chain (or at least the major steps) without direct training. The difference in acquisition

may reflect basic differences in learning rates (or "cognitive" abilities). This analysis suggests that intervention strategies for Subject B should have included a more basic initial skill development program to increase rates of sequencing, learning small chains of behavior, and improving attentional skills.

Procedures that would lend themselves to skill training include precision teaching (White & Haring, 1978) and direct instruction (e.g., Engelmann & Bruner, 1968). These techniques focus on building rates of specific behaviors necessary to perform a particular skill and involve careful monitoring of student progress. Other methods used successfully with behaviorally handicapped children have utilized imitation training. The importance of an imitative repertoire as a basis for acquiring new skills has been demonstrated repeatedly (e.g., Apolloni, Cooke & Cooke, 1977; Peck, Apolloni, Cooke, & Rauer, 1978; Rauer, Cooke, & Apolloni, 1978). In fact, shaping imitation skills is often a starting point for many behavioral programs for severely handicapped populations. Equipped with these methods, teachers or trainers of the handicapped may be able to develop basic skills to a level permitting participation in a reciprocal interaction program.

In interpreting Subject B's performance during intervention, the notion of reciprocity again seems useful. In

this case, the necessary conditions for establishing reciprocity were not met and subsequently, reciprocal interactions did not occur. The definition of reciprocity specifies that equitable rates of reinforcement be present. Although, this condition was structured into the game, the discrepancies in language and other skills between Subject B and his peers may have limited the possibility of achieving equivalence.

Strain, Shores, and Kerr (1976) reported a similar case in a study with behaviorally handicapped children. The subjects who had minimal social contact with their nonhandicapped peers were taught to emit positive social behaviors. This procedure resulted in increased positive social interactions with normal peers. One subject, however, failed to show an increase in social interactions and only emitted the positive social behaviors when directly reinforced by the classroom teacher. The authors noted that this subject was at a much lower developmental level than the rest of the class, including the other handicapped students. The apparent discrepancy in behavioral repertoires was thought to be the major factor contributing to this subject's unchanging behavior. Similarly, in the present study, Subject B was developmentally behind and behaviorally deficient in comparison with his classmates.

Evidence from the reciprocity studies cited and the

present research indicates that reciprocity is an important factor in the development of social behavior. The specific conditions that favor reciprocity, however, have not been determined empirically. It has been suggested that language and developmental skills are crucial factors (Strain & Shores, 1977). If participants have similar skills in these areas the likelihood increases that reciprocal interactions will develop. As mentioned, Subject B had limited spoken language skills in comparison with his peers. The summary of his vocalizations emphasizes the point. Verbal operants in his repertoire included mands, tacts, and one-word intraverbals. It should be noted, though, that more complex forms may have been emitted but if so, they were incomprehensible and thus, were recorded as ambiguous. The observed operants coupled with his low rates of directed vocalizations indicates that Subject B's vocal communication skills were considerably less developed than his peers.

To summarize the clinical findings, the skill development game had differential effects on the two target subjects despite the functional similarity in their handicaps. For Subject A, the intervention contributed to increases in cooperative play, vocalizations, and social interactions. Changes in rates of affective behavior were the most striking results. Subject B, on the other hand, showed opposite

trends. Rates of vocalizations and social interactions remained relatively stable or decreased slightly with the exception of "other" behavior following intervention. In the category of play behavior, the percentage of cooperative play decreased while levels of solitary play rose slightly. These data are accounted for by the notion of reciprocity. In the first case, Subject A met the criteria for reciprocal interactions which are considered to be a major component of social skills development. Subject B, however, did not seem to have necessary language and behavioral skills to establish an equitable reinforcing rate with his nonhandicapped peers.

These data have implications for educators planning mainstreamed programs. They suggest, for example, that social integration will not be achieved necessarily by the act of placing handicapped and nonhandicapped children together for an activity. When pairing children for a reciprocal activity, care should be taken to match children according to their repertoires of relevant skills. The more the participants overlap in areas such as language, the greater the likelihood that reciprocity will be achieved. If repertoires are dissimilar, intensive skill training programs for the handicapped child might be warranted to lessen the differences and thus improve the chances of social interactions. Training methods that might be

appropriate include peer tutoring, imitation training, precision teaching and direct instruction methods.

Methodological Analysis.

The second area of analysis focuses on the appropriateness of the procedures used in the investigation. The issues involved concern whether or not the experimental design was adequate given the goals of the program. Another, equally important question, centers on the generalizability of the intervention. How well the skill development game can be utilized by others in similar settings to facilitate mainstreaming needs to be evaluated. Each of these issues will be discussed in the sections to follow.

Experimental Design. Critical for the evaluation of student behavior change is the availability of individual behavior records. Comparing current performance levels to previously obtained samples enables one to judge the impact of treatment program. The single subject experimental design permits such comparisons. It provides a means of measuring effectiveness across a variety of subjects, settings, therapists, teachers, etc. The multiple baseline across subjects design used in the present study, is useful for the evaluation of new treatment programs when ethical or practical considerations prohibit the use of the more rigorous reversal design (Hersen & Barlow, 1976).

The multiple-baseline across subjects design controls for potential sources of variability by staggering the introduction of the treatment over time for different subjects. Other variables such as special events, time of year, concurrent programs, etc. are less likely to be factors in the treatment's effectiveness. The impact of the intervention is assessed by examining the changes in behavior occurring after the introduction of the treatment. Repeated demonstrations of program effects with many different subjects indicates the reliability of the procedure. Of course, subjects must be selected from a relatively homogeneous population in the initial evaluation of the procedure. Generalization of the methods to different groups can be determined after the procedure has been shown to be effective.

The single subject design proved to be particularly important in the present study due to the variability of subject's entering behaviors. Although the two target children showed a functional similarity in certain social behaviors, the origins of those behaviors were quite different. In interpreting the outcome of the treatment, these differences must be considered. Subject A's repertoire more closely approximated that of the nonhandicapped children in the class than Subject B's. The obtained differences in treatment effects likely reflect this

initial discrepancy. Without the individual behavior records, it is doubtful that treatment effects would have been demonstrated. As mentioned earlier, the results with the two target subjects showed trends in different directions. Had the data been summarized across subjects, no changes would have been evident. The single-subject methodology, then, was crucial in the evaluation of the skill development game's impact on the social behaviors of the handicapped students.

Generalization. The differential effect of the treatment with the two target children indicates limited generality of intervention effects for handicapped preschoolers. Due to the differences in results, a functional relation between the intervention and the behavioral outcome could not be established. Additional subjects would be needed before the general impact of the intervention could be assessed. However, the results suggest that the program may benefit certain children with social behavioral deficits. Judging from Subject B's performance, it seems that such entering skills as attending, following a sequence of steps, and expressive language are needed. Of course, further research would be required with a variety of children in order to draw such a conclusions.

As the program stands, its general applicability to a broad group of children is not likely. Preliminary

findings indicate that there may be some basic skills required for the program to succeed. To improve the range of applicability one might initially incorporate skill training in deficient areas. Once the repertoires were developed to the appropriate level, the intervention could be instituted. In addition to modifying student behaviors prior to intervention, one could also modify the intervention itself to reach a wider group of children. This procedure would require, however, an identification of the critical components in the present program. Once identified, the irrelevant features could be varied according to the needs of the individual. As an example, one might consider altering the academic or pre-academic content of the game. Where it is presently focusing on language and descriptive skills, other topics such as number or color concepts could be substituted. This could be easily accomplished by changing the cards used in the game. Although a component analysis would be necessary to determine which features are critical, one would expect the following list to be among them.

- reciprocity: alternating turns, providing each child with equal roles in the interaction
- language component: game should include verbal interaction between participants
- Reinforcement: provisions for social and tangible

reinforcers to be obtained from each other and an adult supervisor

- Feedback: during the course of the game players should have the opportunity to provide feedback to each other on the accuracy of game play

These components appear to be the essential features of the intervention. The specific materials used including the picture cards, game board, playing pieces, etc. are not likely to be critical features. That is, other similar materials could easily have served the same purpose without affecting the essence of the interaction.

Generalizability across settings and subjects is only one measure of generalization. Drabman (1979) has designed a generalization map which provides a conceptual framework for studying generalization by response class across time, settings, behaviors, and subjects. Each of these areas could be assessed systematically following the application of a treatment program. In the case of the skill development game, the strength of behavior change over time might be an area of particular interest to educators. For example, cooperative play for Subject A showed a steady increase during intervention before levelling off at high percentage of free-play time. These data suggest that her behavior probably had come under control of the natural contingencies present in this

setting. It is likely that this behavior would persist if all other elements remained the same. However, a change in setting, such as placement in a new Kindergarten classroom, might result in a decrease in cooperative play unless programs or activities were added to facilitate integration into the new class. Willems (1977) points out the importance of the setting in controlling behavior. In working with patients in a hospital setting, he notes:

> Many traditional, person-based theories of human behavior assume that independence is largely a matter of individual motivation and thus should reflect a high degree of personal constancy across situations. We find instead that behavioral independence varies dramatically when patients move from one hospital setting to another. (1977, p. 51)

The issue of generalization is a major concern of professionals implementing clinical intervention programs. The single-subject methodology is recommended for monitoring progress of individual clients in the treatment program. In addition to providing the therapist or teacher with a day-to-day record of the client's behavior, the data can be used to make decisions regarding the future course of the treatment (Sidman, 1960; Hersen & Barlow, 1976). To assess fully the degree of generalization would require consideration of both client and intervention factors. The following list summarizes the areas mentioned above

relevant to determining generalizability.

- The client's behavior over a period of time: does the behavior remain stable or decrease?
- The client's behavior in other settings: does the rate of behavior change as the setting changes?
- The effectiveness of the procedure with other subjects from similar populations; does the intervention show consistent effects across a group of subjects with similar handicaps?
- The effectiveness of the procedure with other populations: Can the intervention program be applied to a variety of handicapped populations?
- Is the impact of the intervention affected by modifications of materials or modification of deficient client repertoires: Can other means be used to make the intervention appropriate for children who lack the prerequisite skills?

Ecological Analysis.

The final topic of discussion examines the present research from an ecological perspective. It focuses on the classroom as a whole and looks at the impact of the intervention on all children, not just the target subjects. In addition to studying the changes in the classroom social system, the ecological perspective offers guidelines for

evaluating mainstreaming programs. Although the special education law PL. 94-142 requires yearly evaluations of the educational programs for special needs youngsters, this provision will not necessarily provide data on the success of mainstreaming. Handicapped children have the right to an education in the least restrictive environment and are given individual habilitative plans. In the majority of cases, with the exception of the severely and profoundly handicapped, least restrictive environment will mean integration into a regular education classroom. However, the review of students' progress is generally accomplished by measuring how many objectives outlined in the educational plan were attained. Individual accomplishments are, of course, a primary concern for educators. Obviously, if a student fails to acquire expected skills something may well be wrong with the program or its execution. Yet, restricting the evaluation to student academic progress neglects the effects of integration on the other students in the class. One area of importance in determining success of integration is the interaction between handicapped and nonhandicapped students. If the special needs students are socially and academically segregated from normal peers at the end of the school year, one would hardly conclude that mainstreaming had been effective. It seems essential that mainstreaming be evaluated from the perspective of both

the individual and the larger social system.

A number of the efficacy studies investigating the impact of special education conducted in the last few years have looked at the social adaptability of the retarded child in a public shcool setting (Mercer, 1971; Bruiniks et al., 1974; Goodman et al., 1972). The measures of social adaptability, however, consisted of scores on sociometric questionnaires or peer acceptance scales completed by nonretarded peers. Although the data obtained on the social standing of the handicapped child was informative, it did not offer any insights into how, where, and under what conditions the two groups interacted. Careful observation of the children's behaviors in mainstreamed settings is required, such as the type one would find in a behavior analysis report.

The ecological outlook has recently become a focus of applied behavior analysis research. Interest in this topic was spurred by Edwin Willems (1974) paper "Behavioral Technology and Behavior Ecology" published in the Journal of Applied Behavior Analysis. In it, Willems criticized behavior analysts for not monitoring a variety of behaviors other than the target behavior when researching the effectiveness of intervention programs. The target behavior in Willem's ecological view is simply one behavior among the client's "ecological system" of behaviors. According

to the principles of ecology, a change in one aspect of the system is likely to produce changes in another. Thus, Willems argued that altering the rate of the target behavior probably caused unanticipated changes in other behaviors. He concluded that the impact of a procedure can only be determined if a wide range of behaviors are monitored.

Willem's paper initiated a continuing dialogue between ecologists and behavior analysts so that today, the ecological or environmental perspective occurs with increasing frequency. In 1976, Rogers-Warren and Warren edited a series of papers related to ecology and behavior analysis and noted that at least two ecologies were discussed in the literature. The first type consists of the range of intrapersonal behaviors of the client. The second approach focuses on the client within his physical and social environment. Rogers-Warren and Warren (1976) write "although the second definition is closely aligned with the environmentalists' viewpoint, it is definitely behavioral in perspective. That is, although the environment is viewed as influential, it is considered a potential intervention base. Environmental rearrangement is suggested to support behavior change by working in conjunction with contingency-based interventions" (1976, p. 4). This second view has implications for special education service providers. It suggests that altering classroom environments as well as providing specific

academic programs is an appropriate tactic for mainstreaming programs. Again, it is evident that fully evaluating mainstreaming programs will require information on the relationship between the handicapped and nonhandicapped children within the social system of the classroom.

Several authors have begun to look at the classroom environment as a means of understanding social adaptation. Gottlieb (1978) in a theoretical essay outlines a number of factors which need to be considered and offers an organizational framework for studying handicapped children's social adaptation in school. Among the factors on his list are several obvious variables such as verbal and non-verbal communication skills. Other variables that might be somewhat less apparent were characteristics of the peer group, physical stimuli, characteristics of the teacher, environmental demands such as the nature of the academic tasks and the degree of structure provided. He also stresses the importance of observational research in evaluating social adaptation. However, he points out the observer's limitation in not knowing the historical context in which behaviors are occurring. Extending this reasoning one step further, it appears that repeated observations within the same setting would provide the most information about the interactive patterns in the classroom.

The skill development game used in the present study

is an intervention that could potentially alter both types of ecologies--the behavior of the individual target child and the behavior of peers in the social environment towards the target child. Pairing special needs and non-special needs children together for a reciprocal activity sets up an opportunity for interaction that may or may not occur spontaneously. The nature of the activity is such that participants may become reinforcing agents for each other. Whether or not this occurs, the participants will be acquiring information about each other by virtue of the joint activity. According to the ecological view, if the children become reinforcing agents for each other it is likely that their interaction patterns outside of the game will be affected. Even if reinforcing agent status is not achieved, new information about ones peers can also be viewed as a change in the existing system. In both instances the intervention has the potential to alter the social interaction patterns in the classroom. In the present study, free-play behavior was observed as changes in interaction patterns would be detected most readily during this class period.

As was reviewed in the individual data summaries, the social interaction rates and percentages of cooperative play increased for Subject A. These changes represent both types of ecologies. Not only did Subject A change her own

behavior but the patterns of interactions of the other children were also effected. For example, Subject A interaction with a wider variety of peers in cooperative play after intervention than before. This necessarily indicates that peer behavior was also altered. Further evidence was found in the increase in initiations towards Subject A in social interactions.

Subject B's behavior presents a different trend following intervention but is equally amenable to an ecological interpretation. It will be recalled that unlike Subject A, his percentage of solitary play increased slightly after participation in the skill development game. This unexpected outcome indicates, obviously, that solitary play was more reinforcing than other forms of play, but does not explain why. Several possible explanations can account for these data. For example, it is possible that playing the game served to highlight the differences in skills of the two players. As mentioned, Subject B was at a lower developmental level than his peers. If the discrepancy in skills was made more apparent by the intervention process, peers may have been less inclined to initiate play activities with Subject B. The notion of reciprocity suggests that normal peers found play with other children more reinforcing. The discrepancy that peers may have noted could simply be a difference in response latency or

rates. That is, other children may have exhibited social or game behaviors at a rate more like the normal peer than Subject B. If such a difference existed, it would certainly affect the rate of reinforcement. From the ecological perspective, the change in schedule may have represented a force on the social system which resulted in a change in another part of the system, namely normal peer interactions.

At the same time, one could view these data as an example of an intra-individual ecological change. While the intervention may have altered the normal peer's outlook, it also may have effected Subject B's behavior. It is just as likely that reinforcement rates in the cooperative setting were lower than Subject B experienced in solitary play. Solitary play was obviously very reinforcing as it was a high frequency activity in baseline as well as intervention periods. The shift to higher percentages of solitary play may reflect the change in reinforcement rates or a combination of the rate change and availability of a preferred activity.

From this rudimentary ecological analysis of the impact of the skill development game, it is evident that a complete assessment would be a complex undertaking. At this point in time, the technology for conducting ecological evaluations is not readily available for use in the school

systems. However, useful suggestions for developing this methodology further are available. From the present study, it appears that collecting reliable observational data on a variety of behaviors across many subjects is essential. In addition, it seems that repeated observations are needed to give evaluators or researchers an "historical" context for viewing the data. These recommendations coincide with Willems (1974) assertion that both target and unexpected behaviors should be monitored when conducting a behavior analysis. It should also be noted that accumulating data on the non-special needs students provides a normative sample against which treatment effects can be compared. Walker and Hops (1976) demonstrated the usefulness of normative peer data in evaluating classroom programs for behaivorally disordered children. In a mainstreamed setting, normative data would be useful in measuring the effectiveness of individualized education plans and determining the level of social integration achieved.

Gottlieb (1978) points out the necessity for collecting data on the environmental context including available resources, number of others in the setting, demands placed on subjects by others in control, distribution of others by age and sex, etc. Warren (1976) outlines what he calls "Useful Ecobehavioral Principles". Among them he includes the importance of recording response rates as a basic

datum. He suggests that research be conducted to determine the relation between response rates and therapeutic outcome. Such information would be extremely valuable for educational program planners. For example, if a child's low rate of behavior on a certain task prevents participation with peers in classroom activities, then mainstreaming might be facilitated by including rate building of the deficient skill as an objective in the child's educational program. The rate building technique is currently used as a basis of direct instruction procedures (Englemann & Bruner, 1968). Warren (1976) also discusses the need for researchers to consider generalization issues. Use of Drabman's (1979) "generalization map" would be extremely helpful in outlining generalization measures.

Collecting the types of information outlined above would not only enable one to conduct an ecological analysis but would provide relevant data for analyzing the impact of mainstreaming on the social system of the classroom. Without this information a complete analysis of the effects of mainstreaming cannot be obtained.

Conclusions

In summarizing the research findings, several points regarding the evaluation of mainstreaming programs seem evident. At the level of clinical effectiveness, the data

suggest that a reciprocal intervention such as the skill development game might be a useful tool in promoting social integration of some handicapped students in a mainstreamed classroom. Success of the procedure seems dependent on meeting the basic condition of reciprocity, that is, equitable reinforcement rates among the participants. Factors that appear to affect these rates are competencies in language and basic cognitive skills. If the repertoires of the special needs and non-special needs children are grossly dissimilar, then reciprocity is not likely to develop. As many mainstreamed children present cognitive delays along with motor or other physical handicaps, restructuring the intervention to accommodate the more deficient children is desirable. Increasing the applicability of the procedure to a broader population can be accomplished in at least two ways. One method is to provide preliminary training on the necessary skills for the handicapped child to bring rates of behavior to a level compatible with the nonhandicapped child. A second technique for improving generalizability of the intervention is to modify the intervention itself. Both the subject matter and subject response can be changed to accommodate curriculum needs without altering the basic elements of the game. For example, if a teacher were interested in

teaching number concepts instead of language skills, number cards could be substituted for the action pictures. In addition, teachers could have the players perform a matching response instead of verbally describing the cards. In making modifications, it is suggested that the following features be retained:

- participants have equal roles in the interaction
- participants provide feedback to each other
- participants dispense whatever reinforcers are being used
- verbal praise from supervisory personnel

The second issue regarding evaluation of mainstreaming focuses on the selection of a research methodology. The single-subject design is recommended given the diversity of integrated classrooms in terms of size, composition, resources, grade level, etc. Using repeated measures on a variety of behaviors across both handicapped and nonhandicapped children the impact of mainstreaming can be studied. Using an intervention such as the skill development game offers an opportunity to more directly assess the impact of mainstreaming or to put it more accurately, the effect of integrated interaction.

The single-subject design lends itself to an ecological analysis, the third consideration in evaluating mainstreaming programs. The ecological view is necessary as it allows

for an assessment of the effects of the program on both the handicapped target child and the nonhandicapped student. This seems particularly important when one considers the fact that the nonhandicapped students generally represent the majority of the mainstreamed classroom. Although the technology has not been finely developed for conducting what Warren (1976) describes as an "ecobehavioral analysis", various authors have presented recommendations for ecological evaluations. Included among them are the notions of repeated measures, recording environmental context, use of observational measurement systems, and observing a variety of behaviors in addition to the specific target behavior.

Future Research.

The results of the present study suggest a number of topics for future research. One area concerns the skill development game. A number of replications with a variety of handicapped children is needed to understand under what conditions it promotes behavioral increases in social skills. Another aspect requiring research validation is an analysis of the components of the game. Which features are essential for the game to have an impact needs to be determined empirically. In addition, an assessment of modified versions of the skill development game would pro-

vide useful generalization information.

As mentioned, additional research is needed to develop a technology for conducting ecological analyses in a classroom setting. In addition to refined methods that will undoubtedly require the use of computer facilities, methods are needed which can be used by educators of evaluators on an individual classroom basis who do not have access to computer equipment.

A number of other studies could be generated by testing the effectiveness of initial training on the development of reciprocity. The notion of reciprocity and the factors that determine it is also a topic open to investigation.

Reference Notes

Frankowsky, R. J., & Sulzer-Azaroff, B. Individual and group contingencies and collateral social behaviors. Paper presented at the meeting of the Association for the Advancement of Behavior Therapy, San Francisco, 1975.

Kearney, D., Novak, M., & Olley, G. Interactions of hearing and hearing-impaired 4-year-olds in an integrated preschool. Paper presented at the meeting of the American Psychological Association, New York, August, 1979.

Pertsch, C. F. A comparative study of the progress of subnormal pupils in the grades and in special classes. Unpublished doctoral dissertation, Teachers College, Columbia University, 1936.

Ray, J. S. Behavior of developmentally delayed and nondelayed toddler-age children: An ethological study. Unpublished doctoral dissertation, George Peabody College, 1974.

Reese, E. P. Observing, defining and recording behavior: A first step toward accountability. Workshop presented at the meeting of the Association for Behavior Analysis, Dearborn, Michigan, June, 1979.

Tinjaca, M., & Goetz, E. M. Dual-language reciprocal peer tutoring. Paper presented at the meeting of the American Psychological Association, New York, Aug., 1979.

Yamamoto, J. Y., & Kentschy. An examination of intergrade tutoring experience on attitudinal development of inner dity children. Paper presented at the annual meeting of the California Educational Research Association, San Jose, Calif., November, 1972.

References

Anderson, R. C., & Faust, G. W. Educational Psychology. The Science of Instruction & Learning. Dodd, Mead & Company, 1975.

Apolloni, T., Cooke, S. A., & Cooke, T. P. Establishing a normal peer as a behavioral model for developmentally delayed toddlers. Perceptual and Motor Skills, 1977, 44, 231-241.

Baer, D. M. Perhaps it would be better not to know everything. Journal of Applied Behavior Analysis, 1977, 10, 167-172.

Baldwin, W. D. The social position of the educable mentally retarded in the regular grades in the public schools. Exceptional Children, 1958, 25, 106-108.

Bennett, A. A comparative study of subnormal children in the elementary grades. NY: Teachers College, Columbia University, Bureau of Publications, 1932.

Bergin, A. E., & Strupp, H. H. Changing frontiers in the science of psychotherapy. New York: Aldine-Atherton, 1972.

Blanton, R. L. Historical perspectives on classification of mental retardation. In N. Hobbs (Ed.), Issues in the classification of children. Vol. I. San Francisco: Jossey-Bass, Inc., 1975.

Blatt, B. The physical, personality and academic status of children who are mentally retarded attending special classes as compared with children who are mentally retarded attending regular classes. American Journal of Mental Deficiency, 1958, 62, 810-818.

Bricker, D. D. A rationale for the integration of handicapped and nonhandicapped preschool children. In M. J. Gwalnick (Ed.), Early intervention and the integration of handicapped and nonhandicapped children. Baltimore: University Park Press, 1978.

Bruininks, R. H., Runders, J. E., & Gross, T. C. Social acceptance of mildly retarded pupils in resource rooms and regular classes. American Journal of Mental Deficiency, 1974, 78, 377-383

Buell, J., Stoddard, P., Harris, F. R., & Baer, D. M. Collateral social development accompanying reinforcement of outdoor play in a preschool child. Journal of Applied Behavior Analysis, 1968, 1, 167-174.

Cash, W. M., & Evans, I. M. Training preschool children to modify their retarded siblings' behavior. Journal of Behavior Therapy and Experimental Psychiatry, 1975, 6, 13-16.

Cassidy, V. M., & Stanton, J. E. An investigation of factors involved in the educational placement of mentally retarded children. Columbus: Ohio State University Press, 1959.

Charlesworth, R., & Hartup, W. W. Positive social reinforcement in the nursery school peer group. Child Development, 1967, 38, 993-1002.

Chassan, J. B. Research design in clinical psychology and psychiatry. New York: Appleton-Century-Crofts, 1967.

Chennault, M. Improving the social acceptance of unpopular educable mentally retarded pupils in special classes. American Journal of Mental Deficiency, 1967, 72, 455.

Coleman, T. S., et al. Equality of educational opportunity. Washington, DC: United States Government Printing Office, 1966.

Cook, W. J., & Wollersheim, J. P. The effect of labeling of special education students and the perceptions of contact versus non-contact peers. Journal of Special Education, 1976, 10(2), 187-198.

Cooke, T. P., & Apolloni, T. Developing positive social-emotional behaviors: A study of training and generalization effects. Journal of Applied Behavior Analysis, 1976, 9(1), 65-78.

Dunn, L. M. Special education for the mildly retarded--is much of it justifiable? Exceptional Child, 1968, 35 (1), 5-22.

Drabman, R. S. Generalization Map: A conceptual framework for studying generalization effects. Behavioral Assessment, 1979, 1, 204.

Elenbogen, M. L. A comparative study of some aspects of academic and social adjustment of two groups of mentally retarded children in special classes and regular classes. Dissertation Abstracts, 1957, 17, 2496.

Engelman, S., & Bruner, E. Distar Reading Level I. Chicago: Science Research Associates, 1968.

Feldman, R. S., Devin-Sheehan, L., & Allen, V. L. Children tutoring children: A critical review of research. In V. L. Allen (Ed.), Children as teachers - theory and research on tutoring. New York: Academic Press, 1976.

Gage, N. L., & Berliner, D. C. Educational Psychology. Chicago: Rand McNally College Publishing Co., 1975.

Goodman, H., Gottlieb, J., & Harrison, R. H. Social acceptance of EMR's integrated into a non-graded elementary school. American Journal of Mental Deficiency, 1972, 76, 412-417.

Gottlieb, J. Attitudes toward retarded children: Effects of labeling and academic performance. American Journal of Mental Deficiency, 1974, 79, 268-273.

Gottlieb, J. Attitudes toward retarded children: Effects of labeling and behavioral aggressiveness. Journal of Educational Psychology, 1975, 67, 581.

Gottlieb, J. Observing social adaptation in schools. In G. P. Sackett (Ed.), Observing behavior Volume II: Theory and applications in mental retardation. Baltimore: University Park Press, 1978.

Gottlieb, J., & Budoff, M. Social acceptability of retarded children in nongraded schools differing in architecture. American Journal of Mental Deficiency, 1973, 78, 15.

Guerin, G. R., & Szatlacky, K. Integration programs for the mildly retarded. Exceptional Children, 1974, 41, 173.

Hartmann, D. P. Considerations in the choice of inter-observer reliability estimates. Journal of Applied Behavior Analysis, 1977, 10, 103-116.

Hartup, W. W. Cross-age versus same-age peer interaction: Ethological and cross-cultural perspectives. In V. L. Allen (Ed.), Children as teachers: Theory and research on tutoring. New York: Academic Press, 1976.

Hersen, M., & Barlow, D. H. Single case experimental designs: Strategies for studying behavior change. New York: Pergamon Press, 1976.

Hopkins, B. L., & Hermann, J. A. Evaluating interobserver reliability of interval data. Journal of Applied Behavior Analysis, 1977, 10, 121-126.

Johnson, G. O. Social position of mentally handicapped children in regular grades. American Journal of Mental Deficiency, 1950, 55, 60-89.

Johnson, G. O. Special education for the mentally retarded--a paradox. Exceptional Children, 1962, 29, 62-69.

Johnson, G. O., & Kirk, S. A. Are mentally handicapped children segregated in the regular grades? Journal of Exceptional Children, 1950, 17, 65-68.

Johnson, K. R., & Sulzer-Azaroff, B. The effects of different proctoring systems upon student examination performance and preference. In J. M. Johnston (Ed.), Research and technology in college and university teaching. Gainesville: University of Florida, 1975.

Keller, F. S., & Sherman, T. G. The Keller Plan Handbook. Essays on a personalized system of instruction. Menlo Park, California: W. A. Benjamin, Inc., 1974.

Kratochwill, T. R., & Wetzel, R. J. Observer agreement, credibility, and judgement: Some considerations in presenting observer agreement data. Journal of Applied Behavior Analysis, 1977, 10, 133-140.

Linton, T., Jr. The effects of grade displacement between student tutors and students tutored. Dissertation Abstracts International, 1972, 33, (8-A), 4091.

Long, J., & Madsen, C. H. Five year-olds as behavioral engineers for younger students in a day care center. In E. Ramp and G. Sembs (Eds.), Behavior analysis: Areas of research and application. Englewood Cliffs, NJ: Prentice-Hall, 1975.

MacMillan, D. L. Special education for mildly retarded: Servant or savant? Focus on Exceptional Children, 1971, 2, 1-11.

MacMillan, D. L., Jones, R. L., & Aloia, G. F. The mentally retarded label: A theoretical analysis and review of research. American Journal of Mental Deficiency, 1974, 79, 241-261.

Mercer, T. R. Labeling the mentally retarded. Berkeley, California: University of California Press, 1973.

Meyerowtiz, J. H. Self derogations in young retardates and special class placement. Child Development, 1962, 33, 443-451.

Mohan, M. Peer tutoring as a technique for teaching the unmotivated. Child Study Journal, 1971, 1, 217-225.

Niedermeyer, F. C., & Ellis, P. A. Remedial reading instruction by trained pupil tutors. The Elementary School Journal, 1971, 71, 400-405.

Patterson, G. R., & Reid, J. Reciprocity and coercion: Two facets of social systems. In C. Neuringer and J. Michael (Eds.), Behavior modification in clinical psychology. New York: Appleton-Century-Crofts, 1970. Pp. 133-177.

Payne, R., & Murray, C. Principals' attitudes toward integration of the handicapped. Exceptional Child, 1974, 41, 123.

Peck, C. A., Apolloni, T., Cooke, T. P., & Rauer, S. A. Teaching retarded preschoolers to imitate the free-play behavior of nonretarded classmates: Trained and generalized effects. The Journal of Special Education, 1978, 12, 195-207.

Porter, R. H., Ramsey, B., Trembley, A., Iaccobo, M., & Crawley, S. Social interaction in heterogeneous groups of retarded and normally developing children. An observational study. In G. P. Sackett (Ed.), Observing Behavior. Vol. I. Baltimore: University Park Press, 1978.

Rauer, S. A., Cooke, T. P., & Apolloni, T. Developing nonretarded toddlers as verbal models for retarded classmates. Child Study Journal, 1978, 8(1), 1-9.

Rogers-Warren, A., & Baer, D. M. Correspondence between saying and doing: Teaching children to share and praise. Journal of Applied Behavior Analysis, 1976, 9, 335-354.

Rogers-Warren, A., & Warren, S. F. The developing eco-behavioral psychology. In A. Rogers-Warren and S. F. Warren (Eds.), Ecological perspectives in behavior analysis. Baltimore: University Park Press, 1976.

Rosen, M., Clark, G. R., & Kivitz, M. S. (Eds.). The history of mental retardation. Collected papers. Vol. I. Baltimore: University Park Press, 1976.

Rosenthal, R., & Jacobsen, L. Teachers' expectancies: Determinants of pupils' IQ gains. Psychological Reports, 1966, 19, 115-118.

Seitz, S., & Geske, D. Mothers and graduate trainees' judgements of children: Some effects of labeling. American Journal of Mental Deficiency, 1977, 81 (4), 362-370.

Shotel, J. R., Iano, R. P., & McGettigan, J. F. Teacher attitudes associated with the integration of handicapped children. Exceptional Child, 1972, 38, 677.

Sidman, M. Tactics of scientific research. New York: Basic Books, Inc., 1960.

Skinner, B. F. Why we need teaching machines. Harvard Educational Review, 1961, 31, 377-398.

Skinner, B. F. Verbal behavior. Englewood Cliffs, NJ: Prentice Hall, Inc., 1957.

Snyder, L., Apollini, T., & Cooke, T. P. Integrated settings at the early childhood level: The role of nonretarded peers. Exceptional Children, 1977, 43(5), 262-266.

Soule, D. Teacher bias effects with severely retarded children. American Journal of Mental Deficiency, 1972, 77, 208-211.

Sparks, H. L., & Blackman, L. S. What is special about special education revisited: The mentally retarded. Exceptional Child, 1965, 31, 242-247.

Strain, P. S., Cooke, T. P., & Apolloni, T. Teaching exceptional children: Assessing and modifying social behavior. New York: Academic Press, 1976.

Strain, P. S., & Shores, R. E. Social interaction development among behaviorally handicapped preschool children: Research and educational implications. Psychology in the schools, 1977, 14(4), 493-502.

Strain, P. S., Shores, R. E., & Kerr, M. M. An experimental analysis of "spillover" effects on the social interaction of behaviorally handicapped preschool children. Journal of Applied Behavior Analysis, 1976, 9, 31-40.

Strain, P. S., Shores, R. E., & Timm, M. A. Effects of peer social initiations on the behavior of withdrawn preschool children. Journal of Applied Behavior Analysis, 1977, 10, 289-298.

Walker, H. M., & Hops, H. Use of normative peer data as a standard for evaluating classroom treatment effects. Journal of Applied Behavior Analysis, 1976, 9, 159-168.

Warren, S. F. A useful ecobehavioral perspective for applied behavior analysis. In A. Rogers-Warren and S. F. Warren (Eds.), Ecological perspectives in behavior analysis. Baltimore: University Park Press, 1976.

White, O. R., & Haring, N. G. Exceptional teaching. Columbus, Ohio: Charles E. Merrill Publishing Co., 1976.

White, W. D., & Wolfensberger. The evolution of dehumanization in our institutions. Mental Retardation, 1969, 1, 5-9.

Willems, E. P. Behavioral technology and behavioral ecology. Journal of Applied Behavior Analysis, 1974, 7, 151-165.

Willems, E. P. Steps toward an ecobehavioral technology. In A. Rogers-Warren and S. F. Warren (Eds.), Ecological perspectives in behavior analysis. Baltimore: University Park Press, 1977.

Wright, Judge J. S. Hobson vs Hansen: U.S. Court of Appeals Decision on the District of Columbia's track System. Civil Action No. 82-66. Washington, DC: US Court of Appeals, 1967.

Yelton, A. R., Wildman, B. G., & Erickson, M. T. A probability based formula for calculating interobserver agreement. Journal of Applied Behavior Analysis, 1977, 10, 127-132.

Yoshida, R., & Meyers, C. E. Effects of labeling as educable mentally retarded on teacher's expectancies for change in a student's performance. Journal of Educational Psychology, 1975, 67, 521.

APPENDIX I

Letter to Parents

The Commonwealth of Massachusetts
University of Massachusetts
Amherst 01003

DEPARTMENT OF PSYCHOLOGY

January 18, 1979

Dear Parents:

For the second year, the Early Education Center has agreed to participate in a research program with the Department of Psychology at the University of Massachusetts. The project, under the direction of Dr. Beth Sulzer-Azaroff, will be examining the effectiveness of a "reciprocal peer program" in which special needs children and non-special needs children work together on a language game. More specifically, the program will consist of several five to ten minute periods a week in which pairs of children will play a game designed to enhance communication skills. All play sessions will be supervised continually by the experimenter. Children joining in the activity will receive stickers and small paper cutouts as part of the activity. Occassionally, a special activity may be arranged for the students.

It is hoped that this brief training program will foster good working and social relationships among all the students in the class as well as have an impact on their language skills. To assess the effects of the procedure, measures on both the frequency and types of communication that occur between students will be taken. The majority of the data will be gathered by trained research assistants observing in the classroom. In addition, audio tapes will be made of each training session and several video tapes of free-play periods will be recorded. The taping is seen as an extremely useful way for the research staff to check on the reliability of the collected data. Parents should be assured that all data are considered confidential and will not be used for any other purpose than the study described herein. The data will be shared only with the research team, the preschool staff, and any interested parents. When reporting the results, the names of the children will not be used in order to preserve confidentiality.

If you have any questions concerning the project or the nature of the involvement of the students, please call Leslie Weidenman, a doctoral student at the University, to discuss your concerns. The number is 545-0083. If you are willing to have your child participate, please fill out and sign the enclosed consent form and return it to school as soon as possible.

Thank you.

Leslie Weidenman

Leslie Weidenman

APPENDIX II

Consent Form

CONSENT FORM

Please return this form to the Early Education Center at Fort River Elementary School as soon as possible. Thank you.***

Date: _____

Check one:

_____ I hereby give permission for my son/daughter _____ (Child's name) to participate in the research project planned for the preschool.

_____ I do not wish my son/daughter _____ (Child's name) to participate in the project.

I understand that occassional audio- and/or video-taping may be involved in the data collection process.

This project has been approved by the Amherst School System and the Human Subjects Committee at the University of Massachusetts.

_____
Parent's Signature

APPENDIX III

Daily Percentages of Vocalizations and Social Interactions of Subjects A and B

	BASELINE											INTERVENTION															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
# INTERVALS	48	32	32	32	16	32	32	32	16	32	32	48	32	60	16	64	32	32	40	56	16	48	16	32	48	16	16
VOCALIZATIONS																											
I TO ADULT	.02	0	0	0	0	0	0	.03	0	.06	0	.04	0	0	.06	.11	0	0	0	0	0	0	0	0	0	0	0
I TO PEER	.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.017	0	.02	0	0	.04	0	.19
R TO ADULT	.04	0	.03	0	0	.03	0	0	.06	0	0	0	0	0	.06	.11	0	0	0	0	.06	0	0	0	0	0	0
R TO PEER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.25	0
SMILE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.017	0	.02	.12	.03	0	.06	.06
LAUGH	0	0	0	0	0	0	0	0	0	0	0	.06	0	0	0	0	0	0	0	.089	0	.02	0	0	0	0	.31
SOCIAL INTERACTIONS																											
GIVE ADULT	0	0	0	0	0	.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.06	0	0	0	0
GIVE PEER	0	0	0	0	0	0	0	0	.125	0	0	0	.03	0	0	0	0	0	0	.017	0	0	0	0	0	0	0
RECEIVE ADULT	.05	0	0	0	0	.09	0	0	.06	0	0	0	0	0	0	0	0	0	0	0	0	0	.06	0	0	0	0
RECEIVE PEER	0	0	0	0	0	0	.06	0	0	0	0	.02	.03	0	0	0	0	0	0	.034	0	0	0	0	0	0	0
OFFER ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OFFER PEER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.02	0	0	0	0	0
AFFECT. ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.125	0
AFFECT. PEER	0	0	0	0	0	0	0	0	0	0	0	.05	.08	0	0	0	.25	0	0	0	.12	0	0	0	0	.125	0
AGGRESS. ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AGGRESS. PEER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.03	0	0	0
OTHER ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTHER PEER	.02	.06	0	.06	0	0	0	0	0	0	0	0	0	.03	0	0	.03	0	.125	.178	0	0	0	0	0	.06	.19

SUBJECT A: DAILY PERCENT OF INTERVALS OF VOCALIZATIONS AND SOCIAL INTERACTION BEHAVIORS.

	BASELINE													INTERVENTION										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
# INTERVALS	48	16	16	16	32	16	32	48	32	32	16	32	32	16	32	80	64	16	16	32	48	64	48	16	16
VOCALIZATIONS																									
I TO ADULT	.06	0	.06	.15	0	.19	.03	.04	.06	.06	.06	.03	.03	.25	.03	.01	.06	.06	0	0	.06	.08	0	.06	.06
I TO PEER	.02	0	.06	0	0	0	.06	.08	0	0	.06	.15	.03	.125	0	0	.06	0	.06	0	0	.09	.10	0	0
R TO ADULT	.04	0	.12	.02	.03	.10	0	.02	.09	.06	.19	.03	0	0	0	.08	.01	0	0	0	.04	.06	0	0	0
R TO PEER	0	0	0	0	0	0	0	.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AMBIGUOUS	.23	0	0	.23	.15	0	.25	.19	.09	.19	0	.19	.25	.31	.19	.10	.23	.25	.19	.11	.125	.10	.35	.06	.49
SOCIAL INTERACTIONS																									
GIVE ADULT	0	0	0	0	0	.06	0	.06	0	0	.06	0	.06	0	.08	0	0	0	0	.06	0	.015	0	0	0
GIVE PEER	0	0	0	0	0	0	0	.02	0	0	0	.08	0	0	.03	0	.01	0	0	.03	0	.015	0	0	0
RECEIVE ADULT	0	0	0	0	.03	0	0	.06	0	0	0	0	0	0	0	0	0	0	0	.06	0	.015	0	0	0
RECEIVE PEER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.02	.17	0	0	0
OFFER ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.01	0	0	0	.02	0	0	0	0	0
OFFER PEER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.0	.01	0	0	.06	0	0	0	0	0	0
AFFECT. ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AFFECT. PEER	0	0	0	0	.09	0	.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AGGRESS. ADULT	0	0	0	0	0	.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AGGRESS. PEER	0	0	0	0	.03	0	0	0	0	0	0	0	0	0	0	0	.01	0	0	.03	0	0	0	0	0
OTHER ADULT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTHER PEER	0	0	0	.02	0	0	0	0	0	0	0	.03	0	0	.06	.01	.01	0	0	0	.02	.04	.02	0	0

SUBJECT B: DAILY PERCENTAGE OF INTERVALS OF VOCALIZATIONS AND SOCIAL INTERACTION BEHAVIORS.

APPENDIX IV

Daily Percentages of Vocalizations and Social Interactions of Non-Special Needs Subjects

DAYS		1	2	3	4	5	6	7	8
VOCALIZATIONS									
I:	TO ADULT	.06	0	.21	.12	.4	.40	.06	0
	TO PEER	.12	.125	.19	.06	.25	.12	.25	.167
R:	TO ADULT	.25	0	.06	.25	.09	.19	.125	.1
	TO PEER	.125	.25	.31	0	.25	.125	.25	.2
AMBIGUOUS		.23	0	.06	.12	0	.16	.125	.12
OTHER		0	.19	.16	0	.09	.09	0	0
SOCIAL INTERACTIONS									
GIVE		.06	0	0	0	.06	.08	0	.06
RECEIVE		0	0	0	0	0	.06	0	0
OFFER		0	0	0	0	0	0	0	0
AFFECTION		0	.12	0	0	0	0	0	0
AGGRESSION		0	0	.12	.06	.06	0	0	.06
OTHER		0	0	.06	0	.06	0	0	.06

NON SPECIAL NEEDS STUDENTS: DAILY PERCENTAGE OF INTERVALS OF VOCALIZATIONS AND SOCIAL INTERACTION BEHAVIORS